AD_____

Award Number: DAMD17-00-1-0375

TITLE: Radiographic Techniques and Lesion Detection Performance in Digital Mammography

PRINCIPAL INVESTIGATOR: Walter Huda, Ph.D.
David R. Dance, Ph.D.

CONTRACTING ORGANIZATION: The Research Foundation
of the State University of New York
Albany, New York 12201-0009

REPORT DATE: September 2002

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 2002	Annual (1 Sep 01 - 31 Aug 02)

4. TITLE AND SUBTITLE
Radiographic Techniques and Lesion Detection Performance in Digital Mammography

5. FUNDING NUMBERS
DAMD17-00-1-0375

6. AUTHOR(S)
Walter Huda, Ph.D.
David R. Dance, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Research Foundation of the State University of New York
Albany, New York 12201-0009

E-Mail: hudaw@mail.upstate.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Original contains color plates: All DTIC reproductions will be in black and white.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*
Our experimental work uses a Lorad Digital Mammography system, together with an anthropomorphic breast phantom. We have developed techniques to produce a digital version of masses and calcifications, which will possess an adjustable level of contrast. A 4 Alternate Forced Choice (4-AFC) methodology has been implemented to objectively measure imaging performance, and which has been shown to be markedly superior to subjective methods of assessment of imaging performance. A Monte Carlo code has been developed which enables theoretical calculations to be performed on the expected improvement in observer performance with changing radiographic techniques and phantom thickness. We have performed the following experimental studies: (1) use of an ACR phantom to assess dose & image quality in digital mammography; (2) Observer performance as a function of radiographic technique factors; (3) The importance of size and random noise on lesion detection performance in digital mammography; (4) The effect of lesion location affect lesion detection in digital mammography, which includes both breast thickness and structured (anatomical) background effects. Our results to date indicate that technique factors are of little importance for lesion detection in breast phantoms, but that breast thickness and structured background do affect lesion detection.

14. SUBJECT TERMS
Breast cancer detection; digital mammography; x-ray radiation; x-ray dosimetry; imaging performance

15. NUMBER OF PAGES
102

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Cover.....

SF 298.....

Table of Contents..... 1

Introduction..... 2

Body..... 3

W Huda 3

DR Dance 7

Key Research Accomplishments..... 11

Reportable Outcomes..... 13

Conclusions..... 15

References..... 16

Appendices.....

Appendix 1.Approved Statements of Work A-1

Appendix 2. Research Presentations (8) A-4

Appendix 3. Research Publications (5) A-5

Introduction

This research relates to digital mammography, which separates the process of image acquisition from any subsequent image display. In comparison to screen-film mammography, the amount of radiation used to produce an image could be increased (or decreased) by over an order of magnitude with no significant change on the displayed image intensity. In addition, the quality of the x-ray beam used to acquire the digital radiograph is also a free variable, with no direct relationship to the displayed image contrast. The purpose of our study is to investigate how changes in radiographic technique factors and breast composition affect observer detection performance in digital mammography. Observer performance will be measured as a function of the following four key parameters: (i) radiation exposure (i.e., x-ray mAs value); (ii) x-ray beam quality (i.e., kVp and filtration); (iii) breast thickness; (iv) breast background structure.

This report describes the work accomplished from 1 September 2001 through to 31 August 2002. Appendix 1 lists the approved statement of work as taken from the submitted grant for the PI, sub-contractor. This report describes the work performed by both the Principal Investigator (Dr Walter Huda), and also that of the principal sub-contractor (Dr David Dance).

Digital Mammography system

Full characterization of the Lorad FFDM digital mammography system has been performed. Digital radiographic images have been acquired using this FFDM system, and successfully used in our experiments. We have completed our investigation into the effect of radiographic technique factors on image quality and radiation dose using a conventional ACR phantom. Results obtained in this study quantify the trade-offs between dose and image quality in digital mammography for the detection of simulated masses in an average size breast. Two papers have been published, and these data will serve as a "baseline" against which to compare the results obtained using an anthropomorphic phantom.

Phantom

We have investigated the effect of breast structure. A technique has been proposed to quantify background structure in digital mammograms represented by the anthropomorphic phantom. The method involved taking a simulated lesion and calculating a cross-correlation coefficient, which will measure the extent to which a given region resembles the lesion. This approach will permit a pilot study to be performed that will compare experimental measures of observer performance in areas where the "cross correlation index" is high with other areas where this index takes on a low value. This will be our initial attempt to get direct evidence of the importance of background breast structure.

Workstation

A high quality workstation (Barco Inc.) has been acquired. We have developed software that permits the selection of a region of interest with simulated lesions, adjustable contrast, and optimal image display. Initial experiments indicated that the use of a 9-Alternate Forced Choice (AFC) methodology is not practical, and this has now been replaced by a 4-AFC method. Dr AE Burgess (Harvard University) has assisted us in the practical aspects of implementation of the 4-AFC workstation. The development of 4 4-AFC software package has now been completed and initial results have been published which compare subjective assessment of image quality with an objective metric (92% correct score on a 4-AFC system).

Summary of the key investigations performed to date.

1. Use of the ACR phantom for assessing image quality in digital mammography.

Digital radiographs were obtained of the ACR accreditation phantom as a function of varying technique factors. In one series, the x-ray tube voltage was kept constant (28 kVp) and digital radiographs were generated using tube current-exposure times between 5 and 500 mAs. In a second series, the tube current-time product was kept constant (80 mAs), and the x-ray tube voltage varied between 24 and 34 kVp in 1 kVp increments. Image display on a 5 megapixel diagnostic quality grayscale monitor was optimized for the best display of the simulated fibers, specks and masses that the phantom contains. An additional five images were included at the same technique factors (28 kVp and 80 mAs) to obtain an estimate of the experimental precision. Eight observers indicated the number of objects visible in each image, which were presented in a

random manner. As the mAs increased, the number of fibers visible increased form < 1 at 5 mAs to all being visible at 80 mAs with similar trends observed for the specks and masses. There was a slight increase in object visibility as the x-ray tube voltage increased from 24 to 26 kVp, but performance was essentially constant above this x-ray tube voltage. These results indicate that the ACR phantom is unsatisfactory for use to assess image quality in digital mammography. One problem is that most objects are clearly visible at the normal technique factors used in digital mammography. A second problem is that the effective dynamic range is relatively narrow, as it occupies the "non-clinical" exposure range below 50 mAs where the mean glandular dose is < 1 mGy.

2. Observer performance and radiographic technique factors in digital mammography.

We investigated how changing the kVp and mAs used to acquire digital mammograms affect observer detection performance. A Lorad Full Field Digital Mammography system was used to expose an anthropomorphic breast phantom (the "Rachel" phantom manufactured by RMI) at x-ray tube voltages ranging from 24 to 34 kVp and output factors ranging from 20 to 160 mAs. Digital mammograms were acquired with and without mass lesions at the same exposure parameters. Following a logarithmic transform, subtraction of these two images yielded a digital version of the mass lesion alone. The intensity of the mass was altered by multiplying each pixel by a scaling factor (SF), and the mass was then added to a phantom image taken with no mass lesion. The lesion SF was adjusted to correspond to the observer detection threshold, so that a high SF value corresponds to low lesion conspicuity and vice versa. Lesion visibility was then assessed as a function of the "scaling factor" necessary to visualize the added lesion. For repeat images generated at 28 kVp and 60 mAs, the "scaling factor" at which the lesion became undetectable (i.e., visibility threshold) was 0.62 ± 0.07. At a fixed 60 mAs, increasing the x-ray tube voltage from 24 to 34 kVp resulted in a decrease of "scaling factor" from 0.73 to 0.66. At a fixed 28 kVp, increasing the x-ray beam intensity from 20 to 160 mAs resulted in a reduction of the "intensity reduction factor" from 0.68 to 0.61. This work shows no evidence of any improvement in mass lesion detection in digital mammography with increases in x-ray tube voltage and x-ray beam intensity.

3. Comparison of objective and subjective methods to assess imaging performance in digital mammography.

The assessment of imaging performance in digital mammography is important for selecting the optimum technique factors (kVp/mAs) and for evaluating the utility of new image processing algorithms. We compared the performance of a subjective imaging performance metric with that of an objective method. The subjective method required observers to specify a probability of a lesion being present in a digital mammogram. The objective method used Four Alternate Forced Choice (4AFC) methodology with an observer identifying which one of four images actually contained the lesion. Digital images were obtained of the anthropomorphic breast phantom with/without added lesions as described in (2) at mAs values that ranged from 30 to 120 mAs. The images of the lesions were again extracted, scaled (scaling factor SF) and added back to a phantom image with no mass lesion at varying levels of contrast. In this manner, the detection threshold contrast level was investigated as a function of the mAs used to acquire the digital mammogram, using both the subjective and objective measures of imaging performance. Results obtained in this study showed that the two methods of measuring imaging performance are complementary. Both subjective and objective modes of evaluating image quality have advantages and limitations, and the most appropriate tool will depend on the specific scientific task at hand.

4. Lesion detection in digital mammography.

We investigated how observer detection performance for a mass varies with the lesion location in digital mammography. Digital images were obtained of the anthropomorphic breast phantom with/without added lesions as described in (2). The images of the lesions were again extracted, scaled (scaling factor SF) and added back to a phantom image with no mass lesion at varying levels of contrast. The lesion SF was adjusted to correspond to the observer detection threshold, so that a high SF value corresponds to low lesion conspicuity and vice versa. Six lesion locations were investigated at each of three signal intensities corresponding to high (pixel value (PV) = 13,200), low (PV = 11,900) and average (PV = 12,550) values. Results showed that lesion conspicuity was greater in the low signal regions (SF = 0.300 ± 0.042) than in the high signal regions (SF = 0.379 ± 0.044), with the average signal levels in-between (SF = 0.338 ± 0.046). The average SF for the three signal intensities was 0.343 ± 0.045. The minimum detectable size of a mass lesion was affected by both signal intensity and local structured background pattern, with each of these factors contributing equally to the variability in lesion detection performance.

5. How does lesion location affect lesion detectability in digital mammography.

In this study, we investigated how the thickness of a mass lesion at the observer detection threshold varied with lesion location. Experimental results were compared to a model of how lesion detection might depend on the structured mammogram background. Digital images were obtained of the anthropomorphic breast phantom with/without added lesions as described in (2). The images of the lesions were again extracted. Lesions thus isolated were added at a reduced intensity to a non-lesion digital mammogram with reduced intensities for a 4 Alternate Forced Choice (4-AFC) experiment. The lesion intensity that corresponded to a 92% correct performance level in the 4-AFC experiment was determined ($I_{92\%}$). Values of $I_{92\%}$ were determined at different locations in the anthropomorphic phantom at a range of average intensity values, thereby permitting the importance of structured background on lesion detection to be investigated. The experimental results are being compared with a simple model that may be expected to explain how structured background might affect lesion detection. The model uses a convolution of the lesion and mammogram, with background subtraction. Regions of breast structure at the same scale as the lesion will give a larger signal and the detection task is expected to be more difficult in such regions. Lesion detection ($I_{92\%}$) was found to depend on both the average signal intensity and on the structured background. Mass lesion detection was best in low signal intensity regions (blacks) and was markedly lower in the high signal regions (white).

The average coefficient of variation within a signal intensity region was 12%, comparable to the 10% coefficient of variation between the three signal intensities. The appearance of the 'detectability map' produced by the convolution depends upon the background subtraction, but initial results suggest that a correlation will be found with the values of $I_{92\%}$ from the observer experiments. This study provides empirical evidence as to how lesion location affects detection performance in digital mammography. The empirical detection data are compared with a model for predicting how structured background might affect lesion detection performance. Our initial results indicate that signal intensity and structured background equally affect the detection of mass lesions in mammograms.

6. The importance of size and random noise on lesion detection performance in digital mammography.

We investigated how lesion size and random noise influence lesion detection performance in digital mammography. Digital mammograms were obtained of an anthropomorphic breast phantom with and without simulated mass lesions as described above. Digital versions of the mass lesions, ranging in size from 0.8 to 12 mm, were added back to the breast phantom image. A series of 4 Alternate Forced Choice (4-AFC) experiments were performed to determine the lesion contrast required to achieve a 92% correct lesion detection rate as a function of the lesion size to generate contrast-detail curves. Experiments were performed using identical phantom images (i.e., twinned) as well as using 10 different versions of phantom images obtained using the same techniques but with different random noise patterns. The slope of the contrast detail curves for lesions in anthropomorphic phantom was always positive, indicating that the larger lesions require more contrast for visualization. This behavior contrasts with conventional contrast-detail curves in uniform backgrounds where the slope is generally -0.5. There was no difference observed between twinned experiments and those obtained using different patterns of random noise. Structured anatomical background requires greater contrast for detection of larger lesions, and random noise has negligible effect on low contrast lesion detection.

Sub-contractor (D R Dance)

During the second year of the contract, the principal objectives of the sub-contractor were:

to complete the development of the computer program and voxel phantom for the simulation of the mammographic examination;

to validate the program against measured data, including image noise;

to use the model to calculate detail-signal -to-noise ratios for various simulated abnormalities;

to start the development of simple methods to quantify the nature of local anatomical structure.

To facilitate this, a visit was made to Dr Huda in Syracuse NY, to discuss the computer program development, the results obtained and to make detailed plans for the next stages of the project.

Dr Roger Hunt, a member of Dr Dance's research group, has made important contributions to this work.

Development of the Monte Carlo compute code and breast model

Our existing Monte Carlo program, which simulates the mammographic examination, was written to average over all points in the image plane (Dance *et al*, 2000a, 2000b) and modeled the breast using a simple geometric phantom. For the purpose of this work, it has been necessary to modify the program so that it can calculate measures of image quality (detail signal-to-noise ratio and contrast) at any point in the image for an inhomogeneous phantom. For this purpose the breast is modeled as a large 3-dimensional array of cells (known as voxels).

At the end of the previous year, many of the modifications required to the program had been completed but the changes to allow the transport of photons through the phantom were not complete. This work has now been finished. The Monte Carlo model makes use of the collision density estimator, which in turn requires knowledge of the radiological path length between interaction and image points. It is very time consuming to compute this quantity, as it involves transport through many voxels. An algorithm due to Siddon (1985) has been used, but modified by us (compared with an earlier implementation) to allow increased computation speed. This was achieved by the use of improved sorting/merging methods.

The program has been coded so that each voxel can vary in composition between pure glandular and pure adipose tissues. The voxels can also have other compositions so that the test phantoms used in this project can be simulated, including the ACR phantom (Huda et al., 2002a) and the Rachel phantom (Yaffe et al., 1993). It was noted in last year's report that contact had been established with Prof Yaffe, who had agreed in principle to provide details of the construction of this phantom. Unfortunately, Prof Yaffe subsequently advised us that this information was lost. Since this phantom is predominately constructed from PMMA, we can use instead our own model of the phantom. In the next section it is demonstrated that the computer simulation can predict the image gray level obtained for different thicknesses of PMMA so that to sufficient accuracy, the thickness of PMMA in any row of voxels can be deduced from the gray level in the image. In this way any region of the Rachel phantom can be simulated and calculations of image properties (including signal-to-noise ratios) made.

Validation of the program

An important aspect of the model development is its validation against experimental measurements. For this purpose, exposures have been made for (a) a fixed thickness of a PMMA phantom and tube voltages in the range 25-32 kV and (b) for fixed tube voltage and PMMA phantom thicknesses in the range 3- 6 cm. Regions of interest were selected in each image, and the mean and standard deviation of the gray level in each determined (image noise). The mean gray level and noise were then computed for the same configurations and the results compared. In addition, the contrast of a 4mm PMMA disk superimposed on the ACR phantom was measured and calculated for tube voltages in the range 25-32 kV. Calculated and measured gray levels show good agreement for the variation with both tube voltage and PMMA thickness. The figure below (left) shows calculated and measured values of the image gray level with phantom thickness. The error bars have been estimated based on uncertainties in the PMMA thickness arising from the use of stock sheets. (The good agreement of calculated and measured pixel values with tube voltage and of contrast with tube voltage is shown in Dance et al., 2000). The figure below (right) shows the variation of the calculated and measured values of the image noise for fixed mAs for imaging the ACR accreditation phantom when the tube voltage is varied. The calculated values of image noise have been normalised so that their average value is the same as that for the measurements. (The program makes a relative calculation of noise. This is all that is required to make relative assessment of imaging performance). Finally it is noted that the experimental measurements of noise variation with mAs by Huda et al. (2000b) show that above about 5mAs, the noise behaviour of the Lorad imaging system is consistent with it being quantum limited. In other words, it is not necessary to incorporate other noise sources within the computer model.

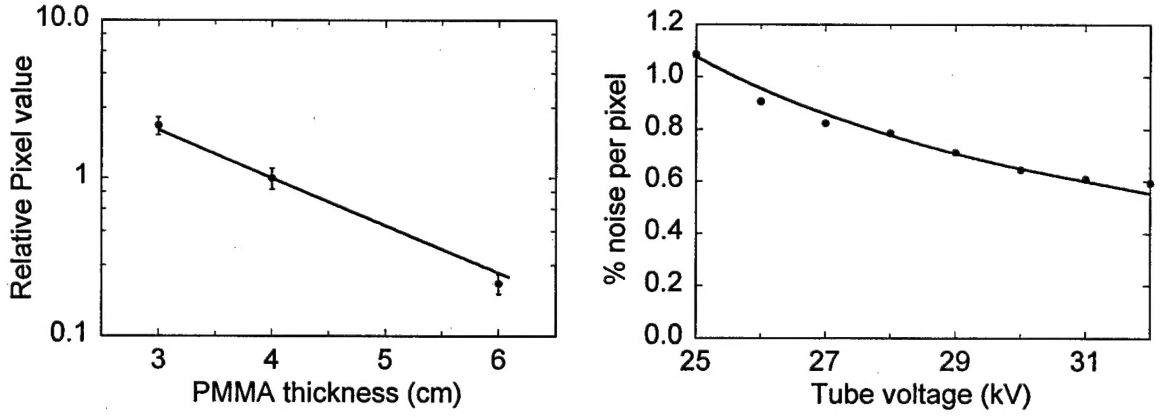


Figure: LEFT: comparison of measured (solid squares) and calculated (solid curve) pixel values with PMMA thickness (solid curve). RIGHT: comparison of measured (solid circles) and calculated (solid curve) values of the % noise per 40 micron pixel at 80 mAs.

Calculations of detail-signal -to-noise ratio for simulated abnormalities; study of the effect of radiographic technique factors

The ACR accreditation phantom has been used in the experimental side of this project and has also been simulated using the computer model. The detail signal-to-noise ratio has been calculated for the simulated abnormalities that this phantom contains. These are 6 nylon filaments in the size range 0.40-1.56 mm (simulating fibrils), 6 sets of aluminium specks in the size range 0.16-0.54 mm (simulating calcifications) and 5 sections of nylon spheres in the thickness range 0.25-2.00 mm (simulating masses). Calculations were done for a range of

radiographic technique factors: tube-current exposure time product in the range 20-160 mAs and of tube voltage in the range 25-32 kV. These parameters were chosen to match the set of experimental measurements with the phantom reported last year (Huda et al. 2002a, b). The results have been compared with and combined with experimental results for the detectability of these objects. The signal-to-noise ratio for the objects just detected experimentally was calculated. The results of this work were presented at the 6th International Workshop on Digital Mammography and will be published in the proceedings of this meeting (Dance et al. 2002). Sample results are shown below.

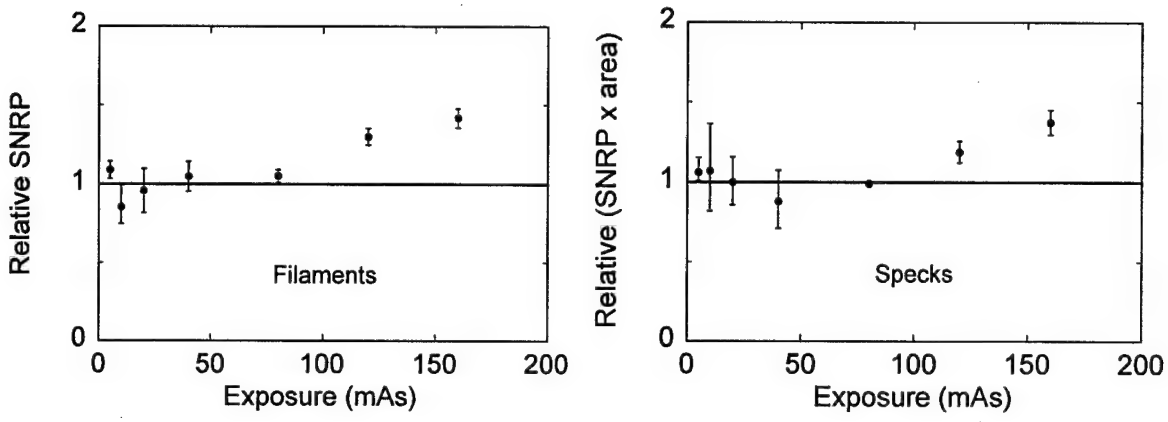


Figure LEFT: Calculated values of SNRP corresponding to the smallest simulated fibril detected experimentally. RIGHT: Calculated values of the detail SNR (SNRP x detail area) corresponding to the smallest simulated calcification detected experimentally. *The error bars represent one standard deviation of the experimental results. Different data points correspond to the detection of details of different sizes.*

The results show that the signal-to-noise ratio per pixel (SNRP) calculated by the model can be used to predict the experimental detectability of simulated fibrils, calcifications and masses against a uniform background. For both filaments and simulated masses and exposures under 80 mAs, the calculated values of the signal-to-noise ratio per pixel for the smallest detail visualised are consistent with a constant value. In other words, objects are detected when the SNRP exceeds a threshold value. Above 80 mAs the size of the objects in the phantom is such that they are above the detection threshold. The results for the simulated calcifications were different. In this case, the product of SNRP and the area of the smallest detectable detail was approximately constant. In other words, for small round objects, the signal-to-noise ratio (SNR) calculated using the total signal (integrated over the image detail - the detail signal-to-noise ratio) is correlated with detectability. These results are entirely consistent with the work reported in Huda et al. (2002a), which suggests that the ACR phantom is not a good test phantom for testing limiting performance of digital mammographic imaging systems. In this respect, the computer model is a good tool for the design of test phantoms and experiments as it can be used to suggest suitable test object sizes. This application of the program is available to the project.

Development of simple methods to quantify the nature of anatomical structure

The task of detecting lesions against a background of normal breast structure is more difficult than that of detecting lesions against a uniform background, which was explored with the ACR phantom. It is anticipated that the detectability will be background dependent and we are investigating methods of quantifying the background structure in a way that may be correlated

with detectability. If such a measure of the structure could be found, this could then be use to predict detectability.

Images of the Rachel phantom and of simulated lesions have been transferred from Syracuse to London and are being analyzed using the IDL image analysis software. We have considered two approaches. In the first approach, the image of the simulated lesion is cross-correlated with the image of normal breast tissue. In regions where the correlation signal is strong it is likely that the background has features which resemble the lesion and the detection task will be more difficult. In the second approach an algorithm designed to detect the abnormality is passed across the image of normal breast tissue. Similar argument would apply about signal strength.

So far we have developed code to explore the first approach. The code is presently being tested and will be used in due course to quantify the structure of the normal tissue for the various simulated lesions used in the multiple-alternative forced-choice (MAFC) experiments with the Rachel phantom.

Key Research Accomplishments - Year 1

Workstation capable of displaying high-resolution digital mammograms has been acquired, tested and configured to operate in the required manner.

Software has been developed that permits the addition of simulated lesions, with a level of contrast that is adjustable by the operator.

Anthropomorphic phantom has been acquired, and prototype lesions that simulate masses and microcalcifications have been manufactured and tested.

The PI and sub-contractor met to plan the appropriate approach to be used for theoretical modeling of the mammography imaging chain; the results of this theoretical modeling will be directly compared with the empirical measures of variations in observer performance.

Access has been obtained to a commercial full field digital mammography system manufactured by Lorad and currently installed at Thomas Jefferson University in Philadelphia.

A non-disclosure agreement has been signed with the manufacturer of the FFDM system (Lorad); this has enabled Lorad to provide proprietary information of the FFDM system which will permit an accurate theoretical model to be developed of how the signal to noise ratio (SNR) changes with technique factors.

Digital images have been acquired of the ACR phantom over the range of mAs and kVp achievable on the FFDM system.

ACR phantom data have been analyzed in terms of image quality (i.e., contrast to noise ratios) and radiation dose; these results provide baseline data with which the anthropomorphic phantom data can be compared.

The optimum x-ray tube voltage for detecting mass type lesion in the ACR phantom was found to be 27 kVp.

An observer study was performed that related detection of lesions (fibers, masses and microcalcification specks) with varying radiographic techniques. These results will be subsequently compared with those that we expect to obtain using an anthropomorphic phantom.

The digital data of the Rachel anthropomorphic phantom have been analyzed in terms of dynamic range and image contrast for the range of radiographic techniques available on the Lorad FFDM system.

A pilot study has been performed studying how mass detection changes with mAs and x-ray tube voltage. The initial results indicated that there was no evident change in performance with either kVp or mAs; the implication of this study is that the structured background is very important and overwhelms the random noise (i.e., quantum mottle) that is present in these images.

Progress has been made in modifications of the Monte Carlo code which will permit theoretical calculations to be performed on the expected improvement in observer performance with changing radiographic techniques.

Key Accomplishments - Year 2

We have used of our digital version of masses to assess observer performance in digital mammography (anthropomorphic phantoms).

We completed the studies with an ACR phantom to assess dose & image quality in digital mammography, which will serve as baseline measurements against which we can compare results obtained using anthropomorphic breast phantoms.

We determined that 9- Alternate Forced Choice experiments were impractical.

A 4-AFC methodology has been implemented to objectively measure imaging performance.

The 4-AFC methodology has been shown to be qualitatively similar to subjective methods of assessment of imaging performance, but with significantly greater experimental precision.

A Monte Carlo code has been developed which enables theoretical calculations to be performed on the expected improvement in observer performance

The Monte Carlo code has been validated by comparison of predictions and experimental measurements.

Calculations with the Monte Carlo code have been combined with observer studies using the ACR phantom to show that with a constant background, lesion detectability can be predicted on the basis of detail signal-to-noise ratio or the signal-to-noise ratio per pixel, depending upon the nature of the lesion.

We have studied how observer performance is affected by radiographic techniques (kV and mAs), and shown that these have little effect on observer detection for most lesions when structured (anatomical) background is present.

We have investigated the importance of breast phantom thickness on lesion detection, and shown that there is a significant drop in detection as thickness increases.

We have commenced experimental work on the importance of local background structure on lesion detection, and shown that this is a significant factor.

We have commenced theoretical studies to try to explain how local background structure may be expected to affect lesion detection.

We have empirically measured lesion detection as a function of lesion size, and shown that contrast detail curve has a positive slope. In other words, larger lesions require more contrast than smaller lesions to be detected.

We have investigated the importance of random noise on lesion detection performance in digital mammography and shown that contrast detail curves are essentially identical for both "twinned" and "random" background structures in 4-AFC experiments.

Reportable Outcomes

Publications

1. AM Sajewicz, W Huda, D Hseuh, KM Ogden, EM Scalzetti, DR Dance. Observer performance and radiographic technique factors in digital mammography. SPIE Medical Imaging 4686 (2002) 107-118.

2. W Huda, AM Sajewciz, KM Ogden, EM Scalzetti, DR Dance. How good is the ACR accreditation phantom for assessing image quality in digital mammography. Acad Radiol 9 (2002) 764-772.

3. DR Dance, R Hunt, AM Sajewicz, W Huda, KM Ogden, M Sandborg, G Alm Carlsson. Comparison of experimental and theoretical assessments of detail visibility in digital mammography. Proceedings of the 6th International Workshop on Digital Mammography (2002) *(In press)*.

4. W Huda, KM Ogden, AM Sajewicz, DR Dance, EM Scalzetti. Comparison of objective and subjective methods to assess imaging performance in digital mammography. Proceedings of the 6th International Workshop on Digital Mammography (2002) *(In press)*.

5. W Huda, AM Sajewicz, KM Ogden, DR Dance. Experimental investigation of the dose and image quality characteristics of a digital mammography imaging system. Accepted by Medical Physics subject to satisfactory revision (September 2002).

Presentations

1. *How good is the ACR accreditation phantom for measuring image quality in digital mammography.* A Sajewicz, W Huda, K Ogden, E Scalzetti. Presented at the *Medical Imaging Perception Society* meeting in Washington DC, September 2001.

2. *Observer performance and radiographic technique factors in digital mammography.* A Sajewicz, W Huda, D Hseuh, K Ogden, E Scalzetti, D Dance. Presented at the *SPIE Medical Imaging* meeting in San Diego CA, February 2002.

3. *Comparison of objective and subjective methods to assess imaging performance in digital mammography.* W Huda, K Ogden, A Sajewicz, D Dance, E Scalzetti. Presented at the 6th *International Workshop on Digital Mammography*, Bremen Germany, June 2002.

4. *Comparison of experimental and theoretical assessments of detail visibility in digital mammography.* D Dance, R Hunt, A Sajewicz, W Huda, K Ogden, M Sandborg, G Alm Carlsson. Presented at the 6th *International Workshop on Digital Mammography*, Bremen Germany, June 2002.

5. *Lesion detection in digital mammography.* K Ogden, W Huda, A Sajewicz, E Scalzetti. Presented at the *American Association of Physicists* in Medicine meeting, Montreal PQ, July 2002.

6. *Dose and image quality in digital mammography.* Presented at the DOD *Era of Hope* meeting, Orlando FL, September 2002.

7. *How does lesion location affect detection performance in digital mammography*? W Huda, K Ogden, E Scalzetti, J Park, R Hunt, D Dance. Submitted to the *SPIE Medical Imaging* meeting in San Diego CA, February 2003.

8. *The importance of size and random noise on lesion detection in digital mammography*. W Huda, K Ogden, E Scalzetti, M Roskopf, D Dance. Submitted to the *European Congress of Radiology*, Vienna Austria, March 2003.

Conclusions

Our experimental work uses a Lorad Digital Mammography system, together with an anthropomorphic breast phantom. We have developed techniques to produce a digital version of masses and calcifications, which will possess an adjustable level of contrast. A 4 Alternate Forced Choice (4-AFC) methodology has been implemented to objectively measure imaging performance, and which has been shown to be markedly superior to subjective methods of assessment of imaging performance. A Monte Carlo code has been developed and validated which enables theoretical calculations to be performed on the expected improvement in observer performance with changing radiographic techniques and phantom thickness. A theoretical approach is being developed which may help to explain how local background structure affects lesion detection. We have investigated the following series of experimental studies: (1) use of an ACR phantom to assess dose & image quality in digital mammography; (2) Observer performance and technique factors; (3) The importance of size and random noise on lesion detection performance in digital mammography; (4) The importance of lesion location affect lesion location in digital mammography, which takes into account variations in breast thickness as well as local structured (anatomical) background. Our results to date indicate that technique factors are of little importance of lesion detection in breast phantoms, but that breast thickness and structured background do affect lesion detection.

References

Dance D R, Thilander Klang A, Sandborg M, Skinner C L, Castellano Smith I A and Alm Carlsson G 2000a. Influence of anode/filter material and tube potential on contrast, signal-to-noise ratio and average absorbed dose in mammography: a Monte Carlo study. Brit. J. Radiol. 73 1056-1067

Dance D R, Skinner C L, Young K C, Beckett J R and Kotre C J 2000b. Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol. Phys Med Biol 45 3225-3240

Dance D R, Hunt R A, Sajewicz A, Huda W, Ogden K M, Sandborg M, Alm Carlsson G 2002. Comparison of experimental and theoretical assessments of detail visibility in digital mammography. Proceedings of 6th International Workshop on Digital mammography, Bremen, Germany, June 2002, in press

Huda W, Sajewicz A M, Ogden K M, Scalzetti E M, Dance D R 2002a. How good is the ACR accreditation phantom for assessing image quality in digital mammography? Academic Radiology 9 764-772

Huda W, Sajewicz A M, Ogden K M, Dance D R 2002b. Experimental investigation of the dose and image quality characteristics of a digital mammography imaging system. Submitted to Med. Phys.

Siddon R L 1985 Fast calculation of the exact radiological path-length for a three dimensional CT array. Med. Phys. 12 252-255

Yaffe M, Byng J W, Caldwell C B, Bennett N R 1993 Anthropomorphic radiological phantoms for mammography. Medical Progress through Technology 19 23-30

Appendix 1.

PI and Sub-contractor statements of work

PI Statement of Work

Task 1 Manufacture of simulated lesions (months 1-9)

- Acquire anthropomorphic phantom/workstation, and configure to import digital mammograms from Lorad system; characterise digital characteristics of the anthropomorphic phantom (months 1-6)
- Acquire tissue equivalent materials and manufacture physical lesions (i.e., size/shape) (months 3-4)
- Develop and verify the production of digital (simulated) lesions techniques (months 4-9)
- Develop method to select region of interest with simulated lesions, adjustable contrast, and optimal image display (months 4-9)

Task 2 Development of 9 Alternate Forced Choice (9AFC) Workstation (months 10-18)

- Develop techniques for determining detection threshold (80% correct score on 9AFC tests) (months 10-12)
- Develop 9AFC software package (months 13-16)
- Validate 9AFC software package (months 16-18)

Task 3 Investigate the effect of radiographic technique factors (months 19-24)

- Acquire digital radiographs with variable kVp/mAs (month 19)
- Perform experiments with variable mAs (exposure) (months 20-21)
- Perform experiments with variable kVp (x-ray beam quality) (months 22-23)
- Analyze experimental data, compare with theoretical expectations (DR Dance subcontract), and write up results (months 23-24)

Task 4 Investigate the effect of breast structure (months 25-36)

- Develop technique for quantifying background structure (DR Dance subcontract) in digital mammograms (anthropomorphic phantom) (months 25-26)
- Acquire digital mammograms for breast structure experiments (month 27) (NB. Guidance on the phantom locations for optimal experiments to by supplied by DR Dance)
- Perform experiments on phantom (breast) thickness and background structure (months 28-30)
- Design and perform additional experiments in light of experimental results obtained to date, to further study the relative importance of "structured background noise" in clinical mammograms (with DR Dance) (31-34)
- Analyze experimental data, compare with theoretical expectations (from Dr DR Dance, subcontractor), and write up results (months 35-36)

Sub-contractor Statement of Work

Task 1 Develop and validate the Monte Carlo model (months 1-12)

- Establish details of the Lorad digital X-ray unit including image receptor and anti-scatter grid and incorporate into Monte Carlo model.
- Match existing tabulations of X-ray spectra to the Lorad spectra.
- Modify current breast model to provide an adequate simulation of anthropomorphic phantom.
- Validate noise estimates using measurements from images of uniform phantoms.
- Extend existing Monte Carlo model as appropriate to improve realism of noise modeling and SNR estimates.

Task 2 Use model to study effect of radiographic technique factors (months 13-27)

- Use model to calculate detail signal-to-noise ratio for various simulated abnormalities as a function of radiation quality and exposure.
- Use calculated results to aid experimental design.
- With PI, compare calculations with experimental results. Write up results.

Task 3 Study of the effect of anatomical structure (months 13-29)

- Develop simple methods of quantifying the local nature of the anatomical structure visualized in the digital images.
- Use these methods to identify regions with similar anatomical structure that can be used for the 4AFC experiment.
- Calculate detail signal-to-noise ratio in these regions and with PI, compare calculations with results of 4AFC experiment
- With PI develop model to explain comparison of experimental and theoretical results.
- With PI, design an experiment to test this model. With PI , compare experimental results with model. Write-up results.

Task 4 Study the effect of breast thickness (months 30-36)

- Use the Monte Carlo model and the quantitative measures of anatomical structure to select five regions of the image with different breast thickness, but with similar structure. These regions will be used in 4AFC experiments by the PI.
- Calculate detail signal-to-noise ratio in these regions and with PI, compare calculations with results of 4AFC experiment. Write-up results

Appendix 2

List of presentations

1. *How good is the ACR accreditation phantom for measuring image quality in digital mammography.* A Sajewicz, W Huda, K Ogden, E Scalzetti. Presented at the *Medical Imaging Perception Society* meeting in Washington DC, September 2001.

2. *Observer performance and radiographic technique factors in digital mammography.* A Sajewicz, W Huda, D Hseuh, K Ogden, E Scalzetti, D Dance. Presented at the *SPIE Medical Imaging* meeting in San Diego CA, February 2002.

3. *Comparison of objective and subjective methods to assess imaging performance in digital mammography.* W Huda, K Ogden, A Sajewicz, D Dance, E Scalzetti. Presented at the 6th *International Workshop on Digital Mammography*, Bremen Germany, June 2002.

4. *Comparison of experimental and theoretical assessments of detail visibility in digital mammography.* D Dance, R Hunt, A Sajewicz, W Huda, K Ogden, M Sandborg, G Alm Carlsson. Presented at the 6th *International Workshop on Digital Mammography*, Bremen Germany, June 2002.

5. *Lesion detection in digital mammography.* K Ogden, W Huda, A Sajewicz, E Scalzetti. Presented at the *American Association of Physicists* in Medicine meeting, Montreal PQ, July 2002.

6. *Dose and image quality in digital mammography.* Presented at the DOD *Era of Hope* meeting, Orlando FL, September 2002.

7. *How does lesion location affect detection performance in digital mammography*? W Huda, K Ogden, E Scalzetti, J Park, R Hunt, D Dance. Submitted to the *SPIE Medical Imaging* meeting in San Diego CA, February 2003.

8. *The importance of size and random noise on lesion detection in digital mammography.* W Huda, K Ogden, E Scalzetti, M Roskopf, D Dance. Submitted to the *European Congress of Radiology*, Vienna Austria, March 2003.

How good is the ACR accreditation phantom for measuring image quality in digital mammography?

I Sajewicz, W Huda, KM Ogden, LM Scalzetti

SUNY Upstate Medical University
Syracuse, NY

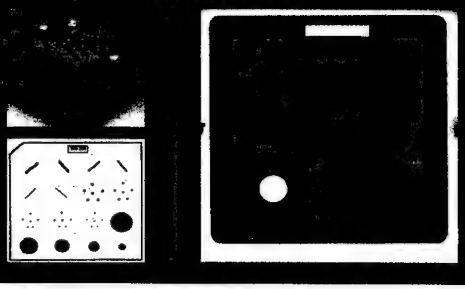
Purpose

Measure how "image quality" in digital mammography changes with kVp mAs using the ACR accreditation phantom

Method

Lorad FFDM

Digital images acquired of an ACR accreditation phantom



At constant 28 kVp

X-ray output varied from 5 to 500 mAs

(12 images)

At constant 80 mAs

X-ray tube potential varied from 24 to 34 kVp

(11 images)

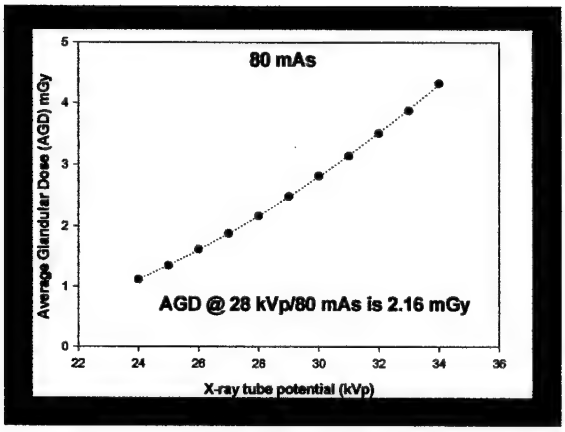
Additional five images generated at

28 kVp and 80 mAs

(i.e. 7 images to measure experimental precision)

Average glandular dose (AGD)
(50% adipose/50% glandular)

Entrance skin exposure
+
kVp/Half Value Layer



ACR phantom images displayed on 5 Megapixel monitor

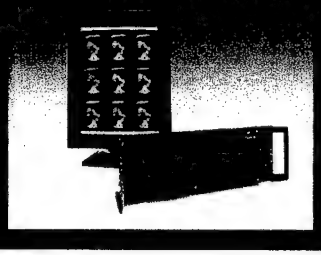
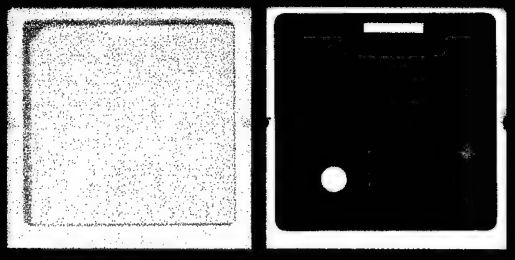
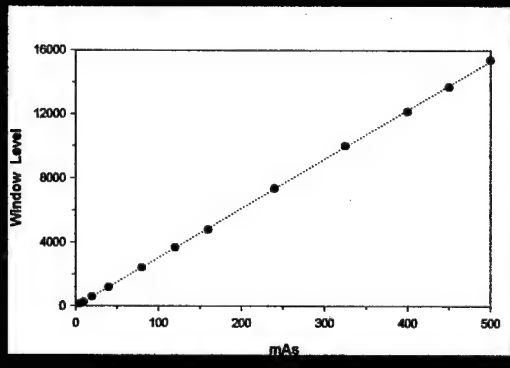
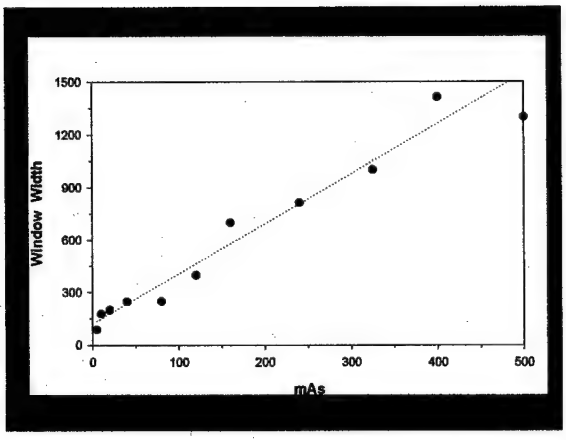


Image display optimized for viewing fibers/specks/masses







Window/Level precision
28 kVp/80 mAs

Window Level 2402 ± 24

Window Width 436 ± 106

Eight (trained) observers

Physicists, RTs, Radiologists

Assessed 28 images in random order in a single session (30 minutes)

visible:

Fibers
(Max 6 & 0.25 scores allowed)

Specks
(Max 30 with 1.0 scores allowed)

Masses
(Max 5 with 0.25 scores allowed)

Results

Inter-observer variability

Mean $\pm \sigma$

Average score for eight observers

Intra-observer variability

σ

Variability of each reader for seven repeats

Fiber
(28 kVp/80 mAs)

Mean: 5.63

Intra-observers σ: 0.31

Inter-observer σ: 0.11

Specks
(28 kVp/80 mAs)

Mean: 23.96

Intra-observers σ: 0.54

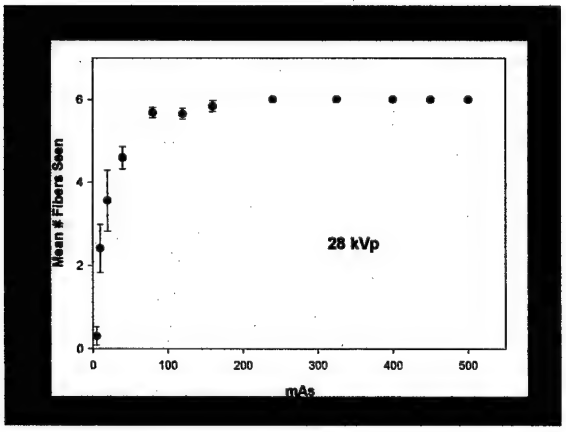
Inter-observer σ: 0.31

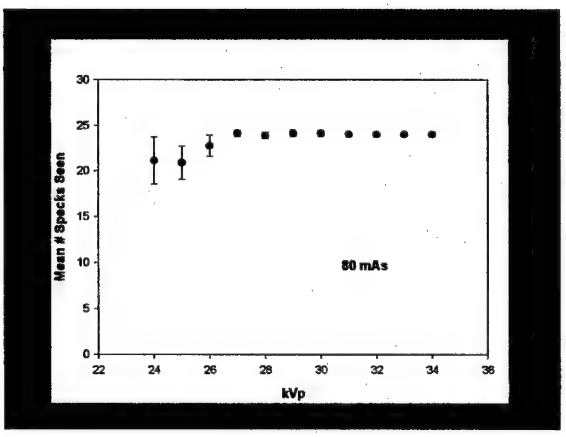
Masses
(28 kVp/80 mAs)

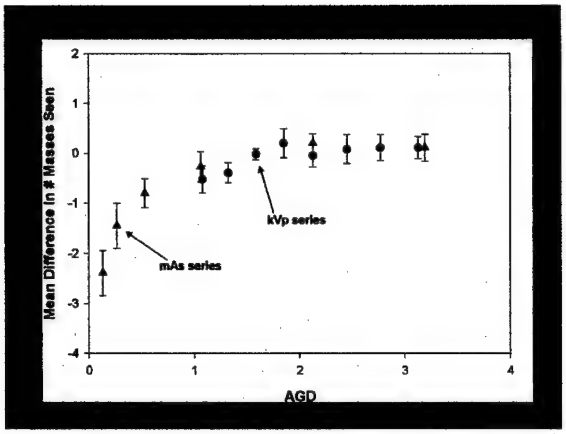
Mean: 3.64

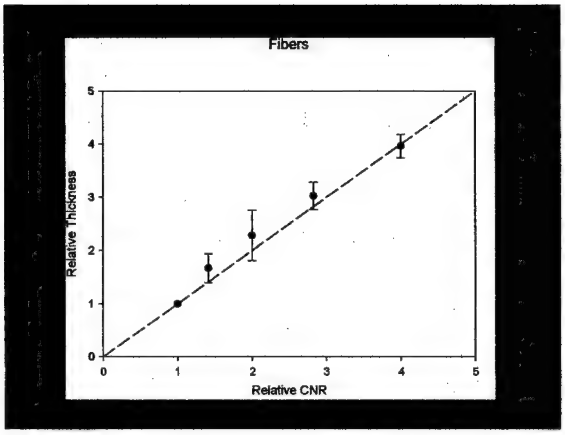
Intra-observers σ: 0.27

Inter-observer σ: 0.43









CONCLUSIONS

For fibers and specks, intra-observer variability > inter-observer variability

For masses, the reverse is true

Increasing the tube current-time product from 5 to 80 mAs

Fibers	0.3 → 5.5
Specks	12 → 24
Masses	1.3 → 4.1

Increasing the tube potential product from 24 to 26 kVp

Fibers	5 → 6
Specks	21 → 24
Masses	3 → 4

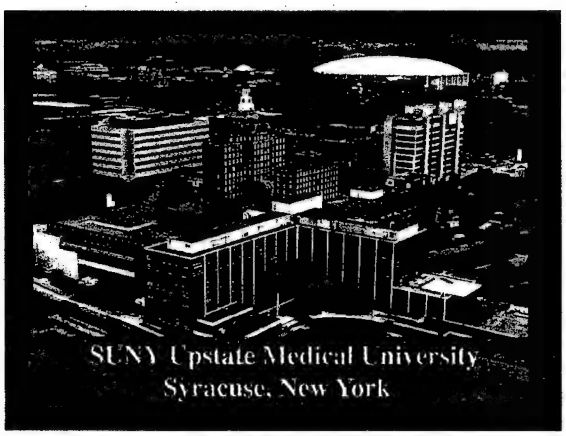
"Dynamic range" of ACR phantom

a average glandular dose < 1.5 mGy

All ACR lesions are visible for AGDs between 1.5 and 3 mGy

Current version of ACR phantom is *not* appropriate for digital mammography

This work was supported in part by a US Army Grant No. DAMD 17-00-1003375



Observer performance and radiographic technique factors in digital mammography.

EM Sajewicz, W Huda, D Hseih, KM Ogden, EM Scalzetti, DR Dance

SUNY Upstate Medical University
Syracuse, NY

Purpose

Investigate how changing kVp & mAs in digital mammography affects lesion detection

Method

Lorad FFDM

Digital image of anthropomorphic phantom

RMI 169
MJ Yaffe

Images acquired without and with lesions → lesions alone

Single calcification (~ 1mm)

Mass lesion (2 x 1.3 cm)

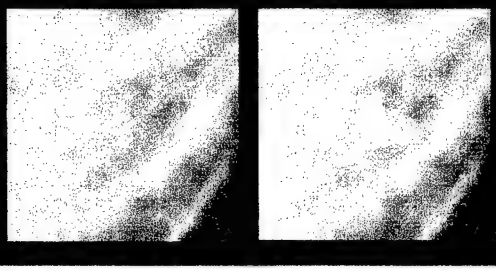
At constant 28 kVp

X-ray tube output:
20, 40, 60, 80, 120 mAs

At constant 60 mAs

X-ray tube voltage:
24, 26, 28, 30, 32 kVp

Lesions digitally added into mammogram regions at 6 different Scaling Factors (SF)



Images at:

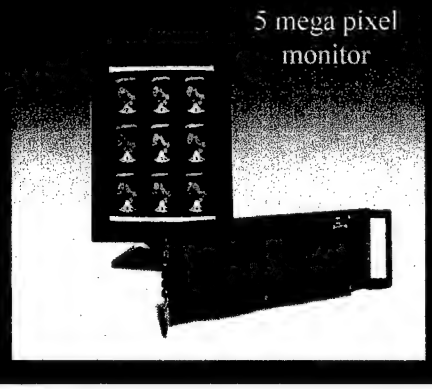
5 different techniques (kVp or mAs)

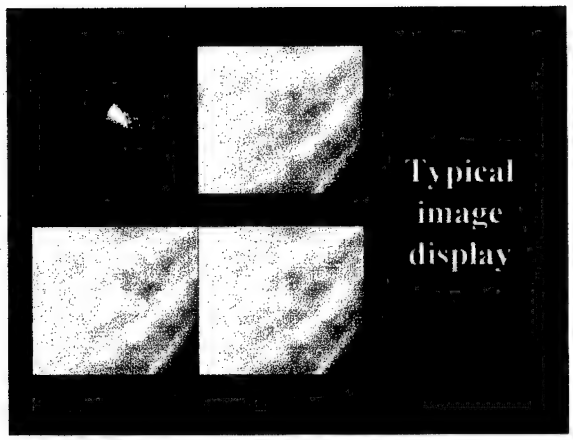
6 scaling factor values

5 repeats
at each SF & technique

Each experiment:

150 randomized images shown to each observer in a single setting





Typical image display

1 or *displayed* region

Window width (min max) 2

Window level max min

further fine adjustment did not improve image display

4 observers

training program

darkened room

no time limit

no additional WL adjustments

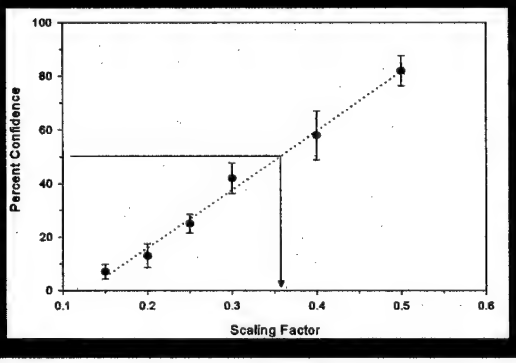
4 Observers score images

probability scale ranging from 0 to 100%

1 or each technique value

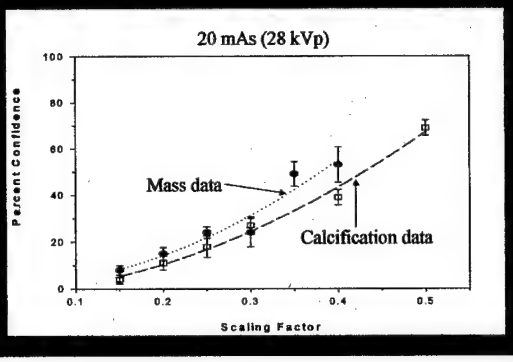
Plot of % confidence (= SI) vs lesion SF

permits the $SF_{50\%}$ to be interpolated

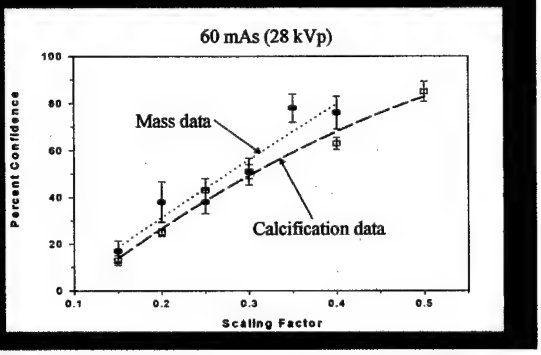


Results

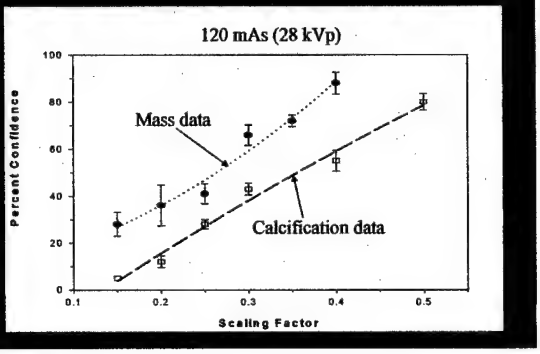
Calc SI = 0.43 Mass SI = 0.38



Calc SI = 0.30 Mass SI = 0.28

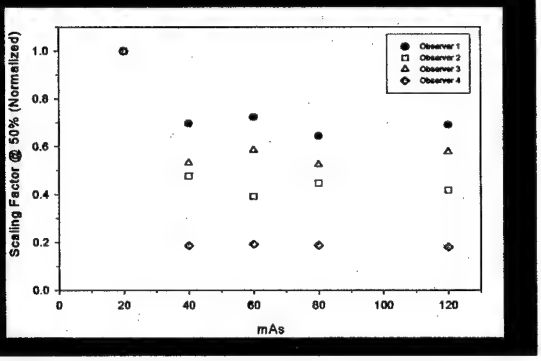


Calc SI = 0.36 Mass SI = 0.26

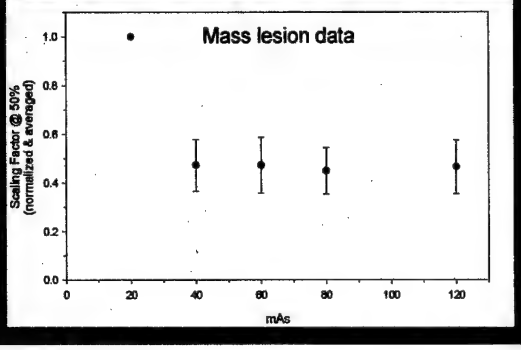


- 2nd order polynomials
- Average r^2 was 0.96 ± 0.02
- shows good internal consistency

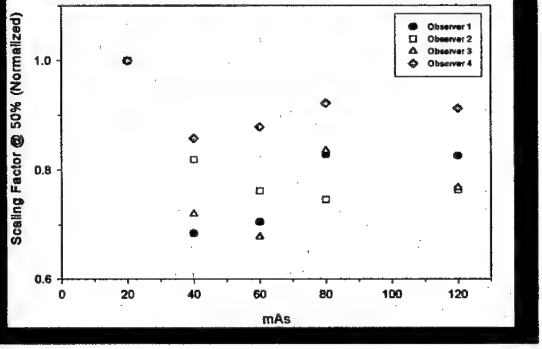
Mass data for mAs



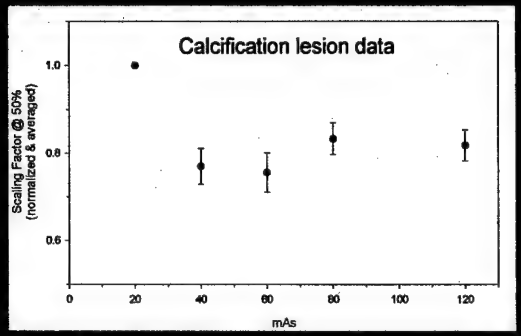
Average score ± standard error



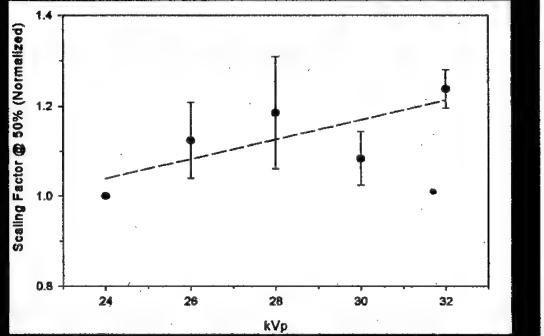
Calcification data for mAs



Average score ± standard error



Calcification data for kVp



Statistical analysis

average slope of 4 readers

vs

slope of zero

t-statistic 2.12

(12 degrees of freedom)

p 0.057

CONCLUSIONS

Method used

Fast & unbiased

Subjective
but consistent results

Qualitative trends
(i.e. pilot study)

For *both* lesions

Lesion visibility was
poorest at 20 mAs

no differences
between 40 – 120 mAs

Trend in kVp data
suggest better detection
at lower kVp values

Need further investigation
(& explanation)

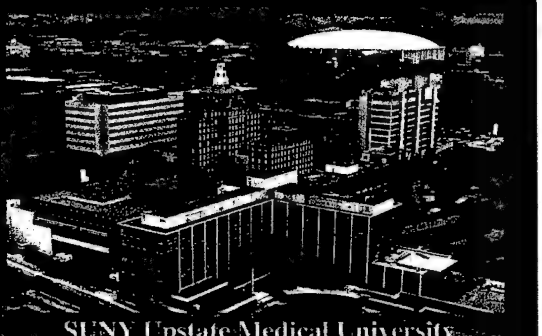
Limitations:

One sample of noise

One sample of background

Large lesions

This work was supported in
part by a US Army Grant
No. DAMD 17-00-1003375



SUNY Upstate Medical University
Syracuse, New York

COMPARISON OF OBJECTIVE AND SUBJECTIVE METHODS TO ASSESS IMAGING PERFORMANCE IN DIGITAL MAMMOGRAPHY

Walter Huda[1], Kent Ogden[1], Anthony Sajewicz[1], David Dance[2], Ernest Scalzetti[1]

[1]Radiology, SUNY Upstate Medical University, Syracuse NY 13210 [2]Joint Department of Physics, The Royal Marsden NHS Trust, London SW3 6JJ

ABSTRACT

Digital images of an anthropomorphic breast phantom were acquired at mAs values between 40 and 120 mAs. Mammograms were obtained with & without added lesions, which permitted the generation of a digital version of the lesion alone which was added back to the phantom image at varying levels of contrast. A subjective method of lesion visibility was compared to an objective method using Four Alternate Forced Choice methodology. Both methods showed observer performance for detecting a mass lesion to be *independent* of x-ray tube output between 40 and 120 mAs.

INTRODUCTION

In this study, we compared a subjective imaging performance approach with an objective method for assessing how lesion detection changes with increasing mAs in digital mammography.

METHOD

Digital image acquisition

Digital images were obtained of an anthropomorphic breast phantom using a Lorad Digital Mammography system at four mAs values (40, 60, 90 and 120 mAs). Digital mammograms were made with/without added lesions, which permitted the generation of a digital version of the lesion alone.

The 1 cm diameter lesion was added back to the digital image of the Rachel phantom alone at varying levels of contrast characterized by a Scaling Factor (SF). Figure 1 (below) shows mammogram background, and the added mass lesion, which were used in both subjective and objective experiments.

Figure 1. Region of a mammogram without (left) and with (right) the added lesion.

Objective observer assessment

The objective method used Four Alternate Forced Choice (4AFC) methodology with an observer identifying which one of four images actually contained the lesion. The results permitted the objective SF to be determined as the value at which the observer accuracy was 92%.

Subjective observer assessment

For the subjective evaluation, six images were produced with lesion visibility ranging from the "extremely difficult" to "easily seen". Five copies of these 24 different images (4 mAs values & 6 SF factors) were generated to produce a series of 120 images.

Three observers specified a probability of a lesion being present on a scale ranging from 0 to 100%. The results permitted the subjective SF value for a 50% probability to be obtained.

RESULTS

Objective results

Figure 2 (left plot) shows the results obtained in the objective 4 AFC experiment, where SF for 92% accuracy is plotted as a function of mAs. The average SF for a 92% correct score was 0.20, and the average coefficient of variation at the four mAs values was 8%.

Subjective results

Figure 2 (right plot) shows the results obtained in the subjective experiment, where SF at 50% observer confidence level is plotted as a function of mAs. The average SF for a confidence score of 50% was 0.46 for this mass lesion, and the average coefficient of variation at the four mAs values was 24%.

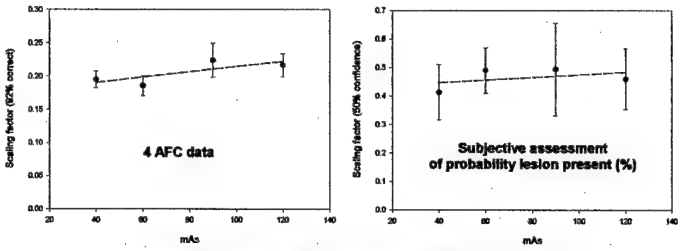


Figure 2. Comparison of objective (4 AFC) and subjective assessments of observer performance.

DISCUSSION

Subjective & objective methods showed similar trends, with no evidence of any improved performance as output increased from 40 to 120 mAs.

The precision (coefficient of variation) in the subjective method was a factor of three larger than the corresponding precision with the 4 AFC approach.

The subjective method had an SF value of 0.46 to achieve a 50% confidence score; a SF of 0.20 resulted in a observers achieving a 92% correct score.

CONCLUSIONS

- Both subjective and objective experiments showed that there was no change in observer performance as the exposure increased from 40 to 120 mAs.
- The subjective method is ideal for performing pilot studies, as it is quick and minimizes systematic errors.
- The objective 4AFC method has a higher precision, and is suited for obtaining definitive scientific data.

ACKNOWLEDGEMENT

This research, under award number DAMD 17-00-1003375, was supported by the Department of Defense Breast Cancer Research Program, which is managed by the US Army Medical Research and Material Command. The authors gratefully acknowledge valuable technical assistance from Dr Zhenxue Jing (Lorad).

Contact person: Walter Huda
Email: hudaw@mail.upstate.edu
Phone (315) 464 7035; FAX (315) 464 7068

Comparison of experimental and theoretical assessments of detail visibility in digital mammography

D R Dance[1], R A Hunt[1], A M Sajewicz[2], W Huda[2], K M Ogden[2], M Sandborg[3] and G Alm Carlsson[3]

[1]The Royal Marsden NHS Trust, London, UK, [2]SUNY Upstate Medical University, Syracuse, NY , USA
[3]Faculty of Health Science, Linköping University, Sweden

Abstract

An experimental and modelling study is being made of the influence of tube voltage, target material and exposure on the performance of digital mammography systems. Digital images of the ACR accreditation phantom at 80 mAs, 25-32 kV and at 28 kV, 5-500 mAs were read by eight observers, and the numbers of fibres, specks and masses visible determined. The computer model simulates photon transport through phantom, anti-scatter grid and image receptor. It calculates image pixels and the signal-to-noise ratio per pixel (SNRP) for the phantom details. For exposures below 100 mAs, the numbers of fibres and masses visualised were found to be consistent with a constant SNRP threshold for detection. For the visualisation of specks, the product of SNRP and speck area was approximately constant. At higher mAs, the number of objects visualised was little influenced by exposure, due to the limited dynamic range of the phantom. The results validate the use of computational models to predict performance for simple detection tasks against a uniform background.

Materials and Methods

An ACR accreditation phantom was exposed using the LoRad digital mammography system and a Mo/Mo spectrum. Images were acquired at 28 kV for exposures between 5-500 mAs and at 80 mAs for tube voltages between 25-32kV.

The phantom (Figure 1 shows an image taken at 28 kV, 80mAs) contains details that simulate fibrils, calcifications and masses (Table 1)

Figure 1

Table 1: Test details embedded in ACR phantom

Detail type	*No*	*Thickness range*
Nylon filament	6	0.40 - 1.56 mm
Alumina speck	6	0.16 - 0..54 mm
Mass	5	0.25 - 2.00 mm

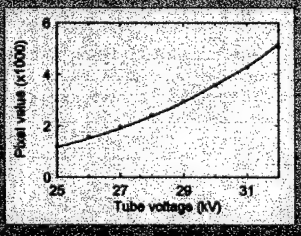
The soft-copy images of the phantom were viewed under controlled conditions by eight readers. The number of objects of each type visible were scored [1].

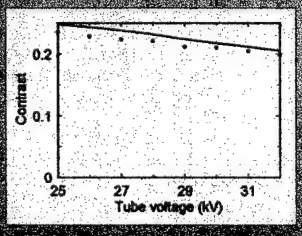
The model used was a Monte Carlo simulation of the mammographic system [2], extended to treat the phantom as an array of voxels. The x-ray spectra were adapted so that they matched measured HVL values. The program could calculate breast dose, entrance air kerma, the energy imparted per image pixel, image contrast and signal-to-noise ratio per pixel (SNRP).

In this way the SNRP was calculated for each detail size and type, exposure and tube voltage.

Validation

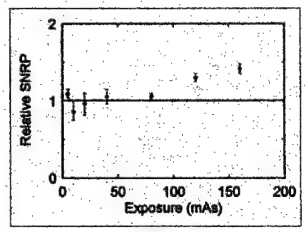
Figure 2 is a validation of the model. It demonstrates the good agreement between calculation and experiment for the variation of pixel value and contrast with kVp.





Results

Figure 3 shows the calculated (normalised) SNRP for the smallest detail detected by the observers at 28 kV for exposures between 5 and 160 mAs. The results are for the detection of nylon filaments (left) and masses (right).



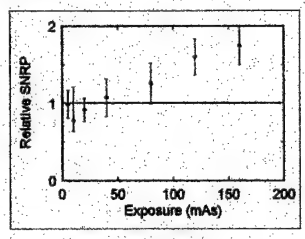


Figure 3: Relative values of the calculated SNRP corresponding to the smallest detectable filament (left) and mass (right) at 28 kV.

The central value of the SNRP at each mAs corresponds to the size of the smallest detectable detail averaged over observers. The error bars represent one standard deviation. Different data points correspond to the detection of objects of different sizes.

In both cases, for exposures of 80 mAs and below, the SNRP for the smallest detail visualised is consistent with a constant value. In other words, **objects are detected when the SNRP exceeds a threshold value.**

Above 80 mAs, some observers see all filaments and the calculated SNRP increases and no longer represents the threshold. A similar comment applies to the masses.

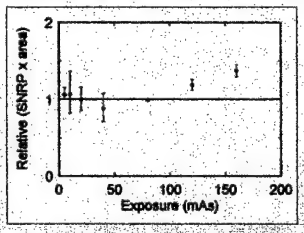


Figure 4: Relative values of (calculated SNRP x detail area) corresponding to the smallest detectable speck at 28 kV.

The results for the specks (Figure 4) are different to those for the filaments and masses. In this case, the objects are small and circular and it was found that for low mAs values, the product of the SNR per pixel and the area for the smallest detectable speck was approximately constant. In other words for small round objects, the SNR calculated using the total signal (integrated over the image of the detail) is correlated with detectability.

The calculations of SNRP for the smallest details detected in the experiments where the kV was varied show similar behaviour to those where the mAs was varied.

Conclusions

The SNRP or SNR calculated by the computer model can be used to predict the experimentally detectability of simulated fibrils, calcifications and masses against a uniform background.

Since the model can also estimate mean glandular dose, it forms a useful tool for the optimisation of the exposure conditions (mAs, kV, target material and filter) so that dose can be minimised for these simple detection tasks. How these tasks relate to the detection of abnormality against a structured background needs to be established

LESION DETECTION IN DIGITAL MAMMOGRAPHY

Kent Ogden, Walter Huda, Anthony Sajewicz, Ernest Scalzetti

Department of Radiology, SUNY Upstate Medical University, Syracuse NY 13210

ABSTRACT

The purpose of this study was to investigate how observer detection performance for a simulated mass varies with the lesion location in digital mammography. A full field digital mammography system was used to acquire images of an anthropomorphic breast phantom with and without a mass lesion. Following a logarithmic transform, subtraction of these two images yielded a digital version of the mass lesion alone. The intensity of the mass was varied by multiplying each pixel by a scaling factor (SF), which was then added to a phantom image with no mass lesion. The lesion SF was adjusted to correspond to the observer detection threshold, so that a high SF value corresponds to low lesion conspicuity and vice versa. Six lesion locations were investigated at each of three signal intensities corresponding to high (pixel value (PV) = 13,200), low (PV = 11,900) and average (PV = 12,550) values. Results showed that lesion conspicuity was greater in the low signal regions (SF = 0.300 ± 0.042) than in the high signal regions (SF = 0.379 ± 0.044). The average coefficient of variation in detection performance at a given signal intensity was 12%, which was comparable to the 10% coefficient of variation between the three signal intensities. These results suggest that signal intensity and structured background are of comparable importance for detection of mass lesions in digital mammography.

METHOD

Image acquisition

A digital mammography system was used to acquire images of an anthropomorphic breast phantom with and without a mass lesion. Following a logarithmic transform, subtraction of these two images yielded a digital version of the mass lesion alone. The intensity of the mass was varied by multiplying each pixel by a scaling factor (SF), which was then added to a phantom image with no mass lesion.

Figure 1 (left) shows a mammogram of the phantom. Regions a, b, and c show typical areas of high, medium, and low pixel intensities, respectively. Six regions were identified at each signal intensity, resulting in a total of 18 regions used in the observer studies. The average high value (white) was 13,200 and the average low value (black) was 11,900.

Observer studies

Figure 2 (below) shows the experimental arrangement used for the observer studies. An 1 cm diameter lesion was added back to the region from the Rachel phantom at varying SF values. Each region from the mammogram is shown to the observer multiple times, with the lesion added in after being multiplied by a scale factor to reduce the lesion contrast.

The observer evaluates the low contrast image using a scale from 0 to 100 - zero means that the observer has absolutely no confidence that the lesion is present in the image, and 100 means that the observer is absolutely certain that the lesion is present. The images are presented with the scale factors varied in a random fashion to minimize observer bias.

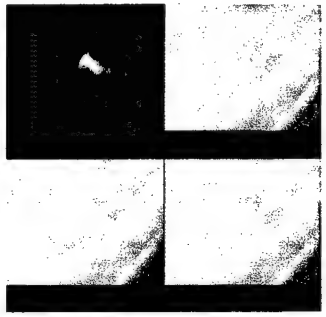


Figure 2 (left) shows a typical presentation of images to observer. The top left image is the lesion reference image, the top right the background reference. The bottom left image shows the lesion in the background at high contrast. The bottom right image contains the background plus lesion at varying contrast levels.

Detection performance

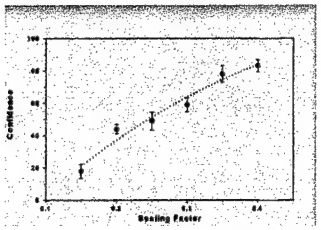


Figure 3 (left) shows results from an observer for a given region. The data show a monotonic increase in the observer score as the SF increases. The SF required to result in a score of 50% was the observer performance metric.

One observer was used to investigate detection performance at each of 18 locations. The experimental precision, based on repeat detection performance measurements, was about 6%. Observer precision was subtracted from measured variability at each intensity level & between the three signal intensities.

RESULTS

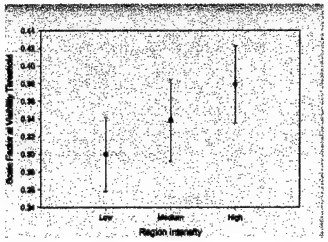


Figure 4 (left) shows how the scale factor required for 50% confidence increased with the region intensity. Lesion conspicuity was greater in the low signal regions (SF = 0.300 ± 0.042) than in the high signal regions (SF = 0.379 ± 0.044), with the average signal levels in-between (SF = 0.338 ± 0.046).

The average coefficient of variation in detection performance at a given signal intensity was 12%, which is comparable to the 10% coefficient of variation between the three signal intensities.

CONCLUSIONS

Mass lesion detection was best in low signal intensity regions (blacks) and was approximately 26% lower in the high signal regions (white).

The average coefficient of variation within a signal intensity region was 12%, comparable to the 10% coefficient of variation between the three signal intensities.

These results show that signal intensity and structured background equally affect the detection of mass lesions in mammograms.

ACKNOWLEDGEMENT

This research, under award number DAMD 17-00-1003375, was supported by the Department of Defense Breast Cancer Research Program, which is managed by the US Army Medical Research and Material Command. The authors gratefully acknowledge valuable technical assistance from Dr Zhenxue Jing (Lorad).

Contact person: Kent Ogden
Email: ogdenk@mail.upstate.edu
Phone (315) 464 5083; FAX (315) 464 5095

DOSE AND IMAGE QUALITY IN DIGITAL MAMMOGRAPHY

Walter Huda PhD, Anthony M. Sajewicz MD, Kent Ogden PhD, Ernest M. Scalzetti MD, David R. Dance,PhD
[1]Department of Radiology, Upstate Medical University, Syracuse, NY; [2]The Royal Marsden NHS Trust, London SW3 6JJ, UK

ABSTRACT

The purpose of this study was to investigate the image quality and absorbed dose characteristics of a commercial digital mammography imaging system, and to identify an optimal x-ray tube voltage for imaging simulated masses. Images were taken of an ACR accreditation phantom using a LORAD digital mammography system. In one experiment, exposures were performed at 80 mAs with x-ray tube voltages varying between 24 and 34 kVp. In a second experiment, the x-ray tube voltage was kept constant at 28 kVp and the technique factor was varied between 5 and 500 mAs. The average glandular dose at each x-ray tube voltage was determined from measurements of entrance skin exposure and x-ray beam half value layer. Image contrast was measured as the digital signal intensity difference for the image of a 4 mm thick acrylic disk. Image noise was obtained from the standard deviation in a uniformly exposed region of interest. The measured digital signal intensity was proportional to the mAs and to the $kVp^{5.8}$. Image contrast was independent of mAs, and dropped by 21% when the x-ray tube voltage increased from 24 to 34 kVp. At a constant x-ray tube voltage, image noise was shown to be approximately proportional to $(mAs)^{(-0.5)}$. At 80 mAs, increasing the x-ray tube voltage from 24 to 34 kVp increased the contrast to noise ratio (CNR) by 78%, and increased the average glandular dose by 285%. At a constant lesion CNR, the lowest average glandular dose value occurred at 27.3 kVp. Imaging simulated masses in a 4.2 cm compressed breast at ~27 kVp results in the lowest average glandular dose.

PURPOSE

To quantify the trade-offs between dose and image quality in digital mammography for detection of mass lesions.

METHOD

Phantom exposures

The full field digital mammography system (LORAD, Danbury CT) had 12 CCDs coupled to a CsI:Tl scintillator.

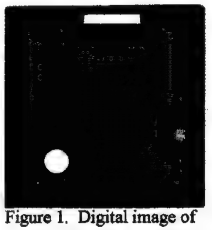


Figure 1. Digital image of ACR accreditation phantom.

Digital images were acquired of an ACR phantom at different values of x-ray tube voltage (kVp) and tube current-exposure time product (mAs). Figure 1 shows a typical phantom image.

Image quality

In one experiment, exposures were performed at 80 mAs with x-ray tube voltages varying between 24 and 34 kVp. In a second experiment, the x-ray tube voltage was kept constant at 28 kVp and the technique factor was varied between 5 and 500 mAs.

Image contrast was measured as the digital signal intensity difference for the image of a 4 mm thick acrylic disk; the corresponding value of noise was obtained from the standard deviation in a uniformly exposed region of interest.

Radiation dose

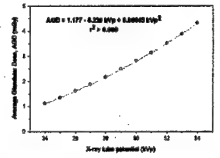


Figure 2. Average glandular dose vs kVp.

The average glandular dose (AGD) at was determined from values of entrance skin exposure and x-ray beam half value layer. Figure 2 shows AGD vs kVp (at 80 mAs).

RESULTS

Signal Intensity

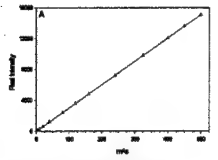


Figure 3. Signal intensity is a linear function of mAs.

Figure 3 shows the expected linear response, with a slope of about 30 pixel values per unit mAs. Extrapolation shows that the system would saturate at a tube current-exposure time product of ~540 mAs (at 28 kVp).

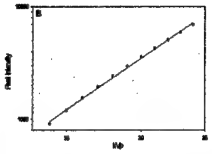


Figure 4. Signal intensity plotted vs kVp on a log scale.

Figure 4 shows the signal intensity as a function of kVp at 80 mAs. The signal obeys a power law relationship. The solid line is a curve fit (slope of 5.80), and signal intensity is proportional to $kVp^{(5.8)}$.

Contrast

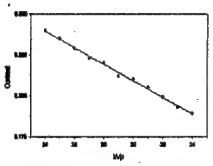


Figure 4. Contrast vs kVp.

As expected, image contrast was independent of mAs. Figure 4 shows image contrast as a function of x-ray tube voltage, which shows that increasing the x-ray tube voltage from 24 to 34 kVp reduced image contrast by 21%.

Noise

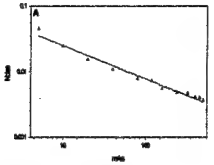


Figure 5. Noise vs mAs.

Figure 5 on the left shows how the image noise varied with mAs, with the solid line a least squares fit ($r^2 = 0.98$) and a slope of -0.506. The slope is very close to the value expected for an imaging system with a noise that is determined by quantum mottle (i.e., slope of -0.500).

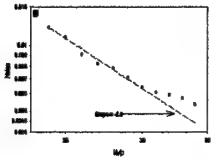


Figure 6. Noise vs kVp at a constant 80 mAs.

Figure 6 shows the measured image noise versus x-ray tube voltage. Increasing the x-ray tube voltage from 24 to 34 kVp reduced the image noise by approximately 55.8%. The corresponding increase in the average glandular dose, however was 285%.

If noise was inversely proportional to signal intensity (i.e., $kVp^{5.8}$), the slope in Figure 6 would be -2.9 (dashed line). Experimental data deviate from this prediction, and the relationship between noise & kVp is *not* simply related to x-ray intensity.

Contrast to noise ratio (CNR)

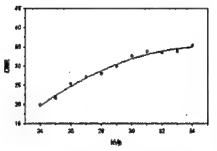


Figure 7. CNR vs kVp at 80 mAs.

Figure 7 shows CNR vs kVp at 80 mAs, which shows that going from 24 to 34 kVp increased the CNR by 78%. At 24 kVp, the CNR increases by 14% per kVp, at 28 kVp the rate of increase was 6.4% per kVp, and at 34 kVp CNR increases by 0.8% per kVp.

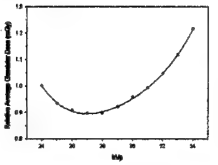


Figure 8. For a constant CNR, the minimum dose is achieved at ~27 kVp.

For a given x-ray tube voltage, the image CNR can be adjusted by modification of the mAs. At a constant lesion CNR, the lowest average glandular dose value occurred at 27.3 kVp as shown in Figure 8. Increasing or decreasing the x-ray tube voltage by 2.3 kVp from the optimum kVp increased the average glandular dose values by 5%.

CONCLUSIONS

X-ray intensity was proportional to $kVp^{5.8}$; image contrast dropped by 21% as x-ray voltage increased from 24 to 34 kVp.

Raising x-ray voltage from 24 to 34 kVp increased CNR by 78% and average glandular dose by 285%.

Imaging simulated masses in a 4.2 cm compressed breast at ~27 kVp results in the lowest average glandular dose.

This work was supported in part by a US Army Grant No. DAMD 17-00-1003375.

1. SUBMIT TO: MI06 Chakraborty/Krupinski

2. IMAGE PERCEPTION, OBSERVER PERFORMANCE, AND TECHNOLOGY ASSESSMENT.

3. How does lesion location affect detection performance in digital mammography?

4. Walter Huda
Department of Radiology, SUNY Upstate Medical University, 750 E Adams Street, Syracuse, NY 13210
Phone (315) 464 7035
FAX (315) 464 7068
Email hudaw@mail.upstate.edu

Kent M Ogden
Department of Radiology, SUNY Upstate Medical University, 750 E Adams Street, Syracuse, NY 13210
Phone (315) 464 5083
FAX (315) 464 7068
Email ogdenk@mail.upstate.edu

Ernest M Scalzetti
Department of Radiology, SUNY Upstate Medical University, 750 E Adams Street, Syracuse, NY 13210
Phone (315) 464 7470
FAX (315) 464 8789
Email scalzete@upstate.edu

Joon Park
Department of Radiology, SUNY Upstate Medical University, 750 E Adams Street, Syracuse, NY 13210
Phone (315) 464 7064
FAX (315) 464 7068
Email parkjk@upstate.edu

Roger Hunt
Department of Physics
The Royal Marsden Hospital
London SW3 6JJ, United Kingdom
Email r.hunt@icr.ac.uk

David R Dance
Department of Physics
The Royal Marsden Hospital
London SW3 6JJ, United Kingdom
Email d.dance@icr.ac.uk

5. Oral or Poster Presentation.

6. SUMMARY.

Purpose. In this study, we investigated how the thickness of a mass lesion at the observer detection threshold varied with lesion location. Experimental results were compared to a model of how lesion detection might depend on the structured mammogram background.

Method. A Lorad Full Field Digital Mammography system was used to expose an anthropomorphic breast phantom. Digital mammograms were acquired with and without mass lesions, thereby permitting a difference image to be generated corresponding to the lesion alone. This isolated lesion was added at a reduced intensity to a non-lesion digital mammogram with reduced intensities using a 4 Alternate Forced Choice (4-AFC) experiment. The lesion intensity that corresponded to a 92% correct performance level in the 4-AFC experiment was determined ($I_{92\%}$). Values of $I_{92\%}$ were determined at different locations in the anthropomorphic phantom at a range of average intensity values, thereby permitting the importance of structured background on lesion detection to be investigated. The experimental results were compared with a simple model that may be expected to explain how structured background might affect lesion detection. The model used a convolution of the lesion and mammogram, with background subtraction. Regions of breast structure at the same scale as the lesion will give a larger signal and the detection task is expected to be more difficult in such regions.

Results. Lesion detection ($I_{92\%}$) was found to depend on both the average signal intensity and on the structured background. Mass lesion detection was best in low signal intensity regions (blacks) and was ~25% lower in the high signal regions (white). The average coefficient of variation within a signal intensity region was 12%, comparable to the 10% coefficient of variation between the three signal intensities. The appearance of the 'detectability map' produced by the convolution was found to depend upon the background subtraction, but correlation was found with the values of $I_{92\%}$ from the observer experiments.

New work. This study provides empirical evidence as to how lesion location affects detection performance in digital mammography. The empirical detection data are compared with a model for predicting how structured background might affect lesion detection performance.

Conclusion. Our initial results suggest that signal intensity and structured background equally affect the detection of mass lesions in mammograms.

7. Keywords. Digital mammography; observer performance; low contrast lesions; structured background;

8. BRIEF BIOGRAPHY. Walter Huda studied Physics at Oxford University and did his Ph.D. in medical physics at the Royal Postgraduate Medical School (University of London). After working at Amersham International, a commercial company specializing

in radioactive products, Dr Huda was at the Manitoba Cancer Treatment and Research Foundation in Winnipeg, Canada and the University of Florida in Gainesville. In 1997, Dr Huda was appointed as a full professor of Radiology at the SUNY Upstate Medical University at Syracuse where his research interests are in digital medical imaging and medical radiation dosimetry.

THE IMPORTANCE OF SIZE AND RANDOM NOISE ON LESION DETECTION IN DIGITAL MAMMOGRAPHY.

Walter Huda[1]
Kent M Ogden[1]
Ernest M Scalzetti[1]
Marsha L Roskopf[1]
David R Dance[2]

[1]**Department of Radiology**
SUNY Upstate Medical University
Syracuse, NY, 13210

[2]**Department of Physics**
The Royal Marsden Hospital
London SW3 6JJ, United Kingdom

Abstract

Purpose. To investigate how lesion size and random noise influence lesion detection performance in digital mammography.

Method. Digital mammograms were obtained of an anthropomorphic breast phantom with and without simulated mass lesions. Digital versions of the mass lesions, ranging in size from 0.8 to 12 mm, were added back to breast phantom image. A series of 4 Alternate Forced Choice (4-AFC) experiments were performed to determine the lesion contrast required to achieve a 92% correct lesion detection rate as a function of the lesion size to generate contrast-detail curves. Experiments were performed using identical phantom images (i.e., twinned) as well as using 10 different versions of phantom images obtained using the same techniques but with different random noise patterns.

Results. The slope of the contrast detail curves for lesions in anthropomorphic phantom was always positive, indicating that the larger lesions require more contrast for visualization. This behavior contrasts with conventional contrast-detail curves in uniform backgrounds where the slope is generally -0.5. There was no difference observed between twinned experiments and those obtained using different patterns of random noise.

Conclusions. Structured anatomical background requires greater contrast for detection of larger lesions, and random noise has negligible effect on low contrast lesion detection.

Appendix 3

List of publications

1. AM Sajewicz, W Huda, D Hseuh, KM Ogden, EM Scalzetti, DR Dance. Observer performance and radiographic technique factors in digital mammography. SPIE Medical Imaging 4686 (2002) 107-118.

2. W Huda, AM Sajewciz, KM Ogden, EM Scalzetti, DR Dance. How good is the ACR accreditation phantom for assessing image quality in digital mammography. Acad Radiol 9 (2002) 764-772.

3. DR Dance, R Hunt, AM Sajewicz, W Huda, KM Ogden, M Sandborg, G Alm Carlsson. Comparison of experimental and theoretical assessments of detail visibility in digital mammography. Proceedings of the 6th International Workshop on Digital Mammography (2002) *(In press)*.

4. W Huda, KM Ogden, AM Sajewicz, DR Dance, EM Scalzetti. Comparison of objective and subjective methods to assess imaging performance in digital mammography. Proceedings of the 6th International Workshop on Digital Mammography (2002) *(In press)*.

5. W Huda, AM Sajewicz, KM Ogden, DR Dance. Experimental investigation of the dose and image quality characteristics of a digital mammography imaging system. Accepted by Medical Physics subject to satisfactory revision (September 2002).

Observer performance and radiographic technique factors in digital mammography.

Anthony M Sajewicz[1], Walter Huda[1*], Denis Hseuh[1], Kent M Ogden[1], Ernest M Scalzetti[1], David R Dance[2]

[1] Department of Radiology, Upstate Medical University, Syracuse, NY 13210
[2] Department of Medical Physics, Royal Marsden Hospital, London, UK

ABSTRACT

In this study, we investigated how changing the kVp and mAs used to acquire digital mammograms affects detection of mammographic lesions. A Lorad Full Field Digital Mammography system was used to expose an anthropomorphic breast phantom at x-ray tube voltages ranging from 24 to 32 kVp and output factors ranging from 20 to 120 mAs. Lesions were added at various intensities to digital mammograms, and lesion visibility was assessed using a subjective probability of the lesion being present, with the image contrast varying from "visible" to "invisible". Four observers ranked the visibility of a large mass lesion (2 cm x 1.3 cm) and a calcification lesion with a diameter of ~1mm. Visibility of both lesions was constant between 40 mAs and 120 mAs (constant 28 kVp), but the visibility of both lesions was significantly lower at 20 mAs. For clinically relevant radiographic techniques, quantum noise does not appear to affect observer performance for detection of lesions in the size range of 1 mm to 2 cm. At a constant mAs, there was a trend showing a reduction in calcification visibility with increasing kVp, but this was not statistically significant ($p = 0.057$).

Key words. Digital mammography; observer performance; radiographic technique factors; low contrast lesions

1. INTRODUCTION

Digital mammography separates the processes of image acquisition from image display, thereby permitting each of these to be individually optimized[1]. Digital mammography offers the operator a choice of adjusting both the x-ray tube voltage (i.e., kVp) and x-ray beam intensity (i.e., mAs) that are not available with analog screen-film imaging systems. The choice of the techniques used to acquire a digital mammogram will generally affect image quality in terms of lesion contrast and image noise (mottle)[2]. Important tasks in digital mammography include the detection of subtle (low contrast) masses and clusters of microcalcifications[3]. In this study, we investigated how changing the kVp and mAs used to acquire digital mammograms affect observer detection performance for the detection of both types of lesions.

2. METHOD

2.1 Digital mammography system

Images were obtained using a Full Field Digital Mammography system (LORAD, Danbury CT), which is a mosaic of twelve 1600 x 1600 pixel Charge Coupled Devices (CCDs) coupled by 2:1 fiber optic tapers to a large area Thallium activated Cesium Iodide (CsI:Tl) scintillator plate. The active image area of the image receptor covers an 18.6 cm x 24.8 cm field. The image pixel matrix size is 4800 by 6400 (40 μm pixels), which corresponds to a Nyquist spatial frequency of 12.5 cycles per millimeter. A conventional linear grid (5:1 grid ratio) is employed to reject scattered x-rays.

Medical Imaging 2002: Image Perception, Observer Performance, and Technology Assessment, Dev P. Chakraborty, Elizabeth A. Krupinski, Editors, Proceedings of SPIE Vol. 4686 (2002) © 2002 SPIE · 1605-7422/02/$15.00

The PC used for image manipulation and display was a DELL® Optiplex computer running at 1000 MHz, and had 512 MB RAM with dual 40 GB hard drives. The video system was a Barco® 5MP1H display system consisting of a 5 million-pixel grayscale monitor and a 10 bit (1024 gray levels) video card. Each pixel has a bit depth of 14 bits and may cover a pixel range from 0 to 16,383. To ensure that images were displayed with the highest possible fidelity, the monitor was calibrated using the BARCO MediCal software and an X-Rite DTP92Q luminance sensor.

2.2 Synthetic images

Images were obtained of an anthropomorphic phantom (Rachel, RMI Madison WI), which simulates a breast as depicted in Figure 1. To study the importance of the amount of radiation used to acquire a digital mammogram, images were obtained at 5 separate mAs values (i.e., 20, 40, 60, 80, & 120 mAs) acquired at a constant x-ray tube voltage of 28 kVp. To study the importance of x-ray beam quality, images were also obtained at five separate kVp values (i.e., 24, 26, 28, 30 & 32 kVp) acquired at a constant 60 mAs.

Fig. 1: Digital radiograph of the anthropomorphic breast phantom (Rachel).

Masses were made from candle wax or soap, which were molded into a variety of shapes including discs, wedges, and spheres. Calcifications were made using crushed calcium carbonate crystals with sizes ranging in diameter from 0.1 mm to 1 mm. For *each* exposure, digital mammograms were acquired with and without mass and calcification lesions, thereby permitting a difference image to be generated corresponding to the lesion alone. Figure 2 shows radiographic images of the three lesions used in these experiments. Lesions, with the background structure subtracted, were selected and used to create the composite experimental images.

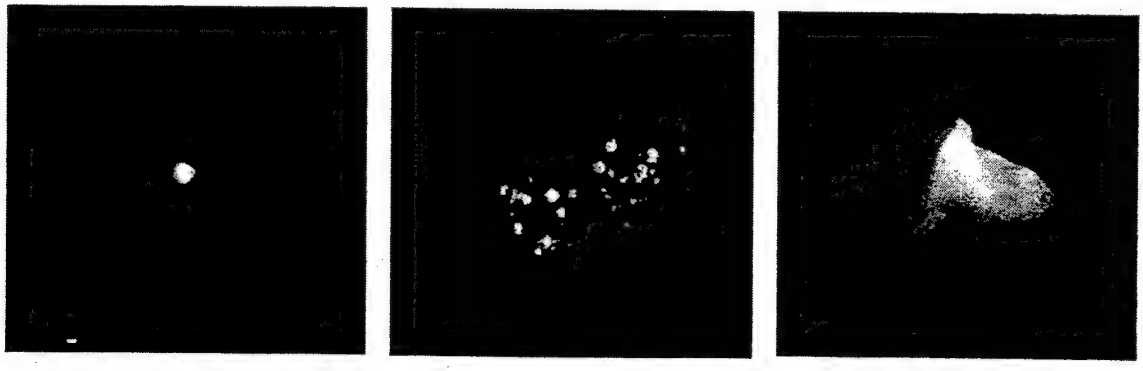


Fig. 2: Single calcification, group of calcifications, irregular shaped mass lesion.

Digital lesions were added to a fixed location in the Rachel phantom obtained at the same techniques. Only the part of the Rachel phantom containing the lesion was used in observer studies, with the size of the displayed region being 4.25 x 4.25 cm (i.e., 850 by 850 pixels). The location of the added lesion was always in the center of the displayed image. The intensity of a given lesion was varied by use of a numerical Scaling Factor (SF), which was used to multiply each pixel value in the lesion. In this manner, a SF value of 0.5 would reduce the lesion contrast by 50%, and a SF value of 2.0 would double the lesion contrast. SF values used in the observer studies were selected on a trial and error basis from pilot studies to determine an average detection threshold. SF values used in this study ranged from 0.15 to 0.4 for the mass lesion and 0.15 to 0.5 for the calcifications.

Table 1 shows the average signal intensities in the lesion location, obtained with and without the lesion being present. These data were obtained by looking at the average intensity value in square/rectangular regions of interest, where the ROI was located in the central region of each lesion location. These data provide information as to the magnitude of the lesion contrast at each specified Scaling Factor (SF), and how these intensity levels varied with the mAs. Table 2 shows the corresponding average signal intensities obtained as a function of kVp for the calcification. Also provided are data on the standard deviation (σ) obtained in a 10 x 10 pixel ROI and a scaling factor of 0.48; the value of σ is an average of the value obtained with & without the lesion. The maximum & minimum pixel intensity values were approximately $\pm$ 2 σ about the mean intensity value. The last two columns in Table 2 list the window and level settings used to display these images.

Table 1. Average intensity level data for variable mAs.

mAs	Mass *		Calcification **	
	Without lesion	With lesion	Without lesion	With lesion
20	561	502	549	493
40	1130	1010	1092	980
60	1699	1517	1655	1487
80	2270	2027	2215	1976
120	3411	3046	3328	2972

* 60 x 120 pixel ROI, SF 0.4
** 10 x 10 pixel ROI, SF 0.5

Table 2. Average intensity level and window width/level data for variable kVp.

kVp	Calcification*		σ	Window width	Window level
	Without lesion	With lesion			
24	704	636	18	1519	1077
26	1198	1077	25	2456	1792
28	1900	1723	25	3590	2756
30	2821	2588	44	5017	3978
32	3999	3685	48	6746	5553

*10 x 10 pixel ROI, SF 0.48

2.3 Observer studies

Each experiment used one of the three lesions shown in Figure 2. For a given radiographic technique (kVp/mAs), the SF was adjusted so that the lesion visibility (contrast) ranged from the "extremely difficult" to "easily seen". A series of such images were generated as a function of either the kVp or the mAs. There were 150 images for a single experiment with five copies of 30 different images (5 radiographic techniques and 6 SF factors). An observer read all 150 images in a single setting, which took 20 minutes on average.

Figure 3 shows the display for a typical experiment, which always showed four separate images. The upper left contains an image of the lesion alone, and the upper right contains an image of the Rachel phantom region of interest (ROI) *without* the lesion. The bottom left is an example of the Rachel phantom ROI with an obvious lesion located at a constant position (center). In the bottom right hand corner, is the sample image that the observer ranks using a probability scale ranging from 0 to 100%. The upper and lower intensity levels of the displayed test image were computed, and used to set the window width and level settings in order to optimize the display. Fine adjustment of the window level and width settings did not improve the image further. Observers participating in the study were not permitted to make any additional modifications to these settings.

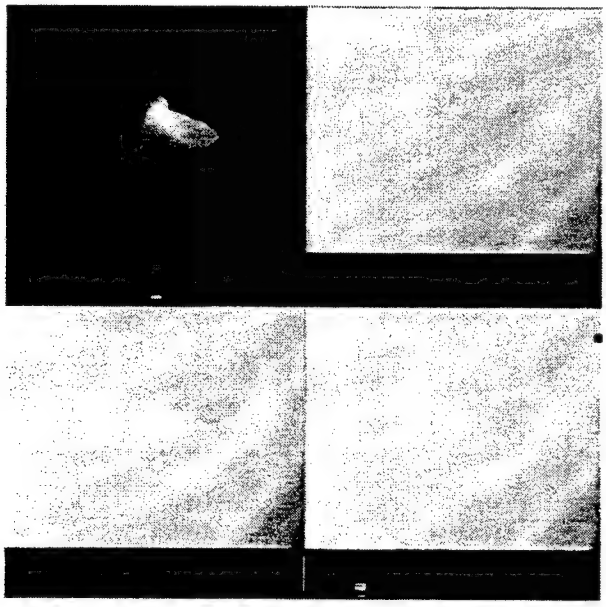


Fig. 3: Experimental display presented to the observer.

Observers underwent a training program to instruct them on the type of lesion being tested, the location of the lesion, and the setup of the display (Figure 3). The training concluded with all observers completing a practice run of about 50 images in order to increase their familiarity and help set their own ranking system for the study. A program was written so that in each experiment, the 150 images were shown to the observers in a random order. The observer sessions all took place in a darkened room with no time limit for viewing each image.

For each lesion and radiographic technique under investigation, the experimental data permit the percent confidence score to be plotted as a function of lesion SF. For a given observer and lesion type, the plot of confidence score (%) against SF permits the SF required to yield a confidence score of 50% ($SF_{50\%}$) to be interpolated. In this manner, we obtained data on the how $SF_{50\%}$ varied with either the mAs or kVp used to acquire the images of the Rachel phantom (+ lesion). Four observers were used to rank the visibility by providing confidence scores of a large mass lesion (2 cm x 1.3 cm) and calcification type lesions with a diameter of ~1mm depicted in Figure 2.

3. RESULTS

3.1 Variable mAs

3.1.1 Observer results

Figure 4 represents one observer's score of the average percent confidence as a function of the scaling factor (SF) over the mAs range investigated. The dotted and dashed lines represent a least squares fit to a second order polynomial for the two sets of data, mass lesion and calcification; the error bars are the computed standard errors for five repeat data sets. The data demonstrate that the observer confidence score (%) increased monotonically with increasing SF value. The fitted curve was used to obtain the $SF_{50\%}$ at each radiographic technique factor used for a given type of lesion. All observers produced results that were qualitatively similar to those shown in Figure 4.

Table 3 shows the individual r^2 averages and their standard deviations for the 4 observers. The data presented here represent the agreement between the observers' confidence scores and the tested scaling factor range. There is no significant difference in the r^2 values among the observers, or among the different types of lesions used in this study. The average value of the coefficient of determination (r^2) for the total of 60 least squares fits was 0.96 ± 0.02.

Table 3. Summary of r^2 values obtained for four observers and three detection tasks.

	Calcification & kVp	Calcification & mAs	Mass lesion & mAs
Observer 1	0.95 ± 0.03	0.99 ± 0.01	0.95 ± 0.02
Observer 2	0.96 ± 0.02	0.98 ± 0.02	0.98 ± 0.01
Observer 3	0.96 ± 0.04	0.98 ± 0.01	0.97 ± 0.02
Observer 4	0.94 ± 0.05	0.97 ± 0.03	0.94 ± 0.06

3.1.2 Absolute results

Table 4 provides the absolute SF values for each reader for the mass lesion, and Table 5 shows the corresponding SF values obtained for the calcification. The results in Table 4 (mass lesion) show marked variability between readers for images obtained at a given mAs; also evident are inter-observer differences in the differential $SF_{50\%}$ score obtained by taking the difference of the scores at 20 mAs from the average scores obtained between 40 and 120 mAs. Although the results for the calcification (Table 5) also show marked inter-observer differences, these are smaller than those observed for the mass (Table 4).

Table 4. Mass lesion average absolute $SF_{50\%}$ scores for 4 observers.

	20 mAs	40 mAs	60 mAs	80 mAs	120 mAs
Observer 1	0.38	0.27	0.28	0.25	0.26
Observer 2	0.75	0.36	0.29	0.34	0.31
Observer 3	0.62	0.33	0.36	0.32	0.36
Observer 4	2.50	0.47	0.48	0.47	0.45

Table 5. Calcification lesion average absolute $SF_{50\%}$ scores for 4 observers.

	20 mAs	40 mAs	60 mAs	80 mAs	120 mAs
Observer 1	0.43	0.29	0.30	0.36	0.36
Observer 2	0.49	0.40	0.38	0.37	0.38
Observer 3	0.31	0.22	0.21	0.26	0.24
Observer 4	0.42	0.36	0.37	0.39	0.39

3.1.3 Relative results

Figure 5 shows the computed $SF_{50\%}$ for the large mass lesion, where the data have been normalized to unity at 20 mAs for each observer. Visibility of the mass lesion was essentially constant between 40 and 120 mAs (constant 28 kVp), but significantly lower at 20 mAs. There is a marked variability between the four observers in the difference between the 20 mAs value and the average values obtained between 40 and 120 mAs. It was notable that for three of the four observers at 20 mAs, the plot of confidence score vs SF required an *extrapolation* to yield the $SF_{50\%}$ value; this procedure was not necessary for readings between 40 and 120 mAs data, as these $SF_{50\%}$ values were obtained by *interpolation*. Figure 6 shows the $SF_{50\%}$ for the calcification lesion. Qualitatively, the observer study results are similar for both types of lesions studied, with significantly inferior lesion visibility evident at (only) the 20 mAs technique value.

Table 6 summarizes the experimental data when averaged for all four observers in scoring the mass and calcification lesions while varying the mAs; the standard deviation data in Table 6 correspond to the standard errors computed for the four observers. The data in Table 6 demonstrate that the value of $SF_{50\%}$ at 20 mAs is significantly higher than at the other four mAs values, and that there is no significant difference in $SF_{50\%}$ between 40 and 120 mAs for either the mass or the calcification. It is notable that use of 20 mAs to acquire the Rachel image had a greater impact on the large mass lesion than for the calcification. Reducing the x-ray tube output to 20 mAs required the contrast for the mass lesion to be increased by a factor of 2.2; the corresponding increase in image contrast for the calcification lesion was markedly lower (i.e., 1.3).

Table 6. Relative $SF_{50\%}$ results averaged over four observers.

mAs	Mass lesion	Calcification lesion
20	1.0	1.0
40	0.47 ± 0.11	0.77 ± 0.04
60	0.47 ± 0.12	0.76 ± 0.05
80	0.45 ± 0.10	0.83 ± 0.04
120	0.47 ± 0.11	0.82 ± 0.04

3.2 Variable kVp

Figure 7 is a graph of the average $SF_{50\%}$ of the four observers as a function of the kVp. The error bars represent the standard error for the observers. The dashed line is a least squares fit to a best line, with a coefficient of determination (r^2) of 0.56. At a constant mAs, there was a trend showing a reduction in calcification visibility with increasing kVp.

A statistical analysis was performed of the measured slopes of the SF vs kVp curves for each individual reader, and whether the average slope for the four readers differed from a curve with a slope of zero. The five data points for each reader represent five degrees of freedom, two of which are used up in the regression estimate of the linear curve fit (Figure 7). With four readers, we have 12 degrees of freedom, and the resultant t-statistic was 2.12 which corresponds to a p value of 0.057. Although these results are not statistically significant, our findings clearly suggest that it would be of considerable interest to perform additional experiments to clarify the issue of whether modification of kVp affects observer performance.

4. DISCUSSION

The Full Field Digital Mammography system used in this study is a quantum noise limited imaging system[2] and the (random) noise level may be taken to be proportional to $mAs^{-0.5}$. Increasing the x-ray tube output from 20 to 120 mAs will therefore reduce the level of mottle by a factor of $6^{0.5}$, or 2.45. The results in Figures 5 and 6 clearly show that lesion visibility is *not* quantum noise limited since there was no significant difference in lesion visibility between 40 and 120 mAs. Both the mass and calcification lesions showed similar behavior despite their being over an order of magnitude difference in lesion size. This result is surprising because one would expect quantum noise to be of greater importance for small lesions rather than a large lesion. The sharp drop in observer performance at 20 mAs suggests that random noise becomes more important than structured background at these low exposure levels. It is important to note that an average breast requires ~ 100 mAs to produce a typical clinical mammogram, and 20 mAs is therefore *much* lower than is likely to be encountered clinically.

In general, higher voltages will reduce image contrast and also reduce the amount of image noise; the effect on the resultant contrast to noise ratio is therefore problematical and will depend on the specific conditions involved. Evidence in the medical imaging literature has generally shown that image CNR improves with increasing voltage (at constant mAs) in head and body CT[4,5]. In digital mammography performed with an ACR accreditation phantom, image CNR has also been shown to increase with kVp (at constant mAs)[2]. These physical measurements are in agreement with observer studies with images of an ACR mammography accreditation phantom generated at different kVps & constant mAs[6]. The results depicted in Figure 7, however, suggest that observer performance was *lower* at the highest kVp. One possible reason for this might be the fact that the dynamic range for a real breast (or anthropomorphic phantom) image can limit the displayed contrast of a given lesion. In addition, our experiments included structured backgrounds, which have been shown to produce contrast-detail curves with a positive slope[7]. If structure limits detection, then quantum noise reduction as kVp increases will be unimportant[8].

The method used to measure observer performance in this study is based on a subjective evaluation made by the observer. Problems associated with this type of subjective evaluation are well known and include inter- and intra-observer variability[9,10]. In this study, we took special precautions to minimize these difficulties. Each observer was extensively trained using a large number of test images, and the training was generally terminated when each observer felt confident that he could maintain a consistent scoring criterion. When the test image was viewed, three other images were displayed at the same time (lesion alone; mammogram alone; obvious lesion in mammogram), and the lesion location was know by the observer, and was fixed. Of greater importance is the fact that a series of randomized images containing the complete range of techniques (kVp or mAs) and the range of SF being investigated produced 30 test images. Repeating these 30 images a total of five times permitted one observer to read a set of 150 randomized images in about 20 minutes.

The results shown in Figure 4 and the corresponding r^2 data summarized in Table 3 suggest that this type of experiment has good internal consistency, and is evidently capable of producing reliable data showing *relative* trends rather than *absolute* performance levels. Our approach cannot make quantitative comparisons with theoretical predictions of how observer performance should change with technique factors. For example, doubling the radiation in a quantum noise limited imaging system will increase the CNR by 41.4%, but it would be unrealistic to expect this to be matched by a subjective observer confidence score. The major limitation of our study is the fact that we only used a *single* sample of the random noise at each technique value. The experimental protocol described in this work is relatively easy to perform, and is expected to serve a valuable role (i.e., pilot study) *prior* to performing more involved forms of analysis. Observer studies such as Receiver Operating Characteristic (ROC) analysis[11,12] or M-Alternate Forced Choice (AFC) methodology[13] are more definitive, but also require a great deal of effort to execute.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. Zhenxue Jing Ph D for assistance with the experimental work and useful discussions on digital mammography. LORAD provided access to the digital mammography imaging system. Paul Sheehe is thanked for help with the statistical analysis, and we also thank Arthur Burgess for helpful discussions. This work was supported in part by a US Army Grant No. DAMD 17-00-1003375.

REFERENCES

[1] M. J. Yaffe, "Digital mammography," *RSNA Categorical Course in Breast Imaging*, pp. 229 – 238, 1999.

[2] W. Huda, A. M. Sajewicz, K. Ogden, and D. D. Dance, "Experimental investigation of the dose and image quality characteristics of a digital mammography imaging system," Submitted to *Med. Phys.*, 2001.

[3] C. J. Vyborny, and R. A. Schmidt, "Mammography as a radiographic examination: an overview," *RadioGraphics* **9**(4), pp. 723 - 764, 1989.

[4] W. Huda, E. M. Scalzetti, and G. Levin, "Technique factors and image quality as functions of patient weight at abdominal CT," *Radiology* **217**, pp. 430 – 435, 2000.

[5] W. Huda, J. Chang, K. A. Lieberman, and M. L. Roskopf, "Radiation dose and image quality in head CT examinations," *RSNA Scientific Program*, 258PH-p, pp. 160, 2001.

[6] W. Huda, A. M. Sajewicz, K. Ogden, E. M. Scalzetti, and D. D. Dance, "How good is the ACR accreditation phantom for measuring image quality in digital mammography?" Submitted to Academic Radiology, 2001.

[7] A. Burgess, F. L. Jacobson, and P. F. Judy, "Human observer detection experiments with mammograms and power-law noise," *Med. Phys.* **28**(4), pp. 419 – 438, 2001.

[8] E. Samei, M.J. Flynn, and W.R. Eyler, "Detection of subtle lung nodules: relative influence of quantum and anatomic noise on chest radiographs," *Radiology* **213** pp. 727 - 734, 1999.

[9] L. Loo, K. Doi, M. Ishida, C. Metz, H. Chan, Y. Hagshida, and Y. Kodera, "An empirical investigation of variability in contrast-detail diagram measurements," *SPIE*, 419, pp. 68 - 76, 1983.

[10] G. Cohen, D. L. McDaniel, and L. K. Wagner, "Analysis of variations in contrast-detail experiments," *Med. Phys.* **11**(4), pp. 469 – 474, 1984.

[11] C. Metz, "ROC methodology in radiologic imaging," *Invest. Radiol.* **21**, pp. 720 – 733, 1986.

[12] D. D. Dorfman, K. S. Berbaum, and C. E. Metz, "Receiver operating characteristic rating analysis. Generalization to the population of readers and patients with the jackknife method," *Invest. Radiol.* **27**(9), pp. 723 – 31, 1992.

[13] A. Burgess, "Comparison of receiver operating characteristic and forced choice observer performance measurement methods," *Med Phys* **22**(5), pp. 643 – 656, 1995.

Corresponding author hudaw@mail.upstate.edu; Phone (315) 464 7035; FAX (315) 464 7068. Mailing address: Radiation Physics, University Hospital, 750 E Adams Street, Syracuse, NY 13210 USA.

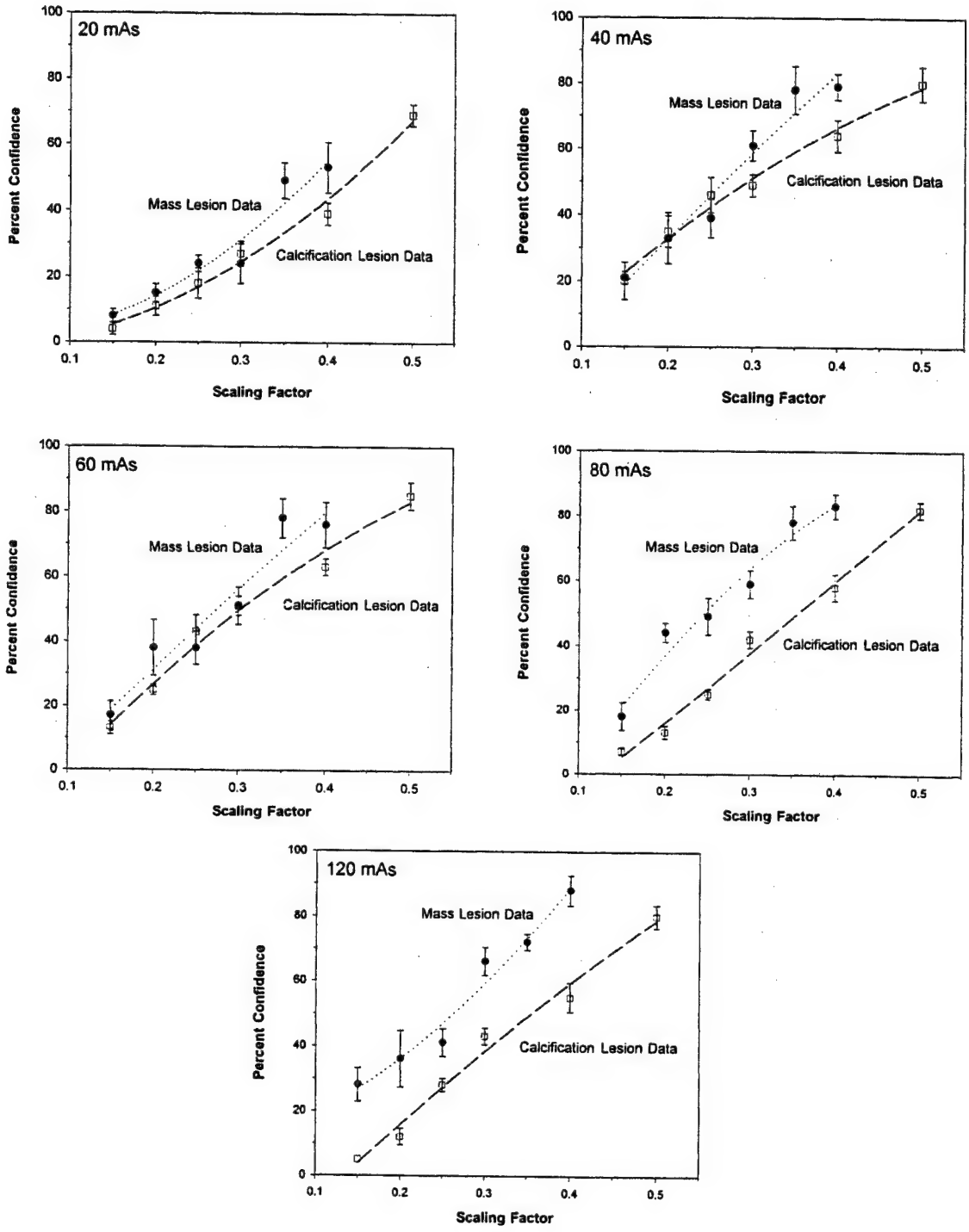


Fig. 4: Average percent confidence as a function of the Scaling Factor for one observer.

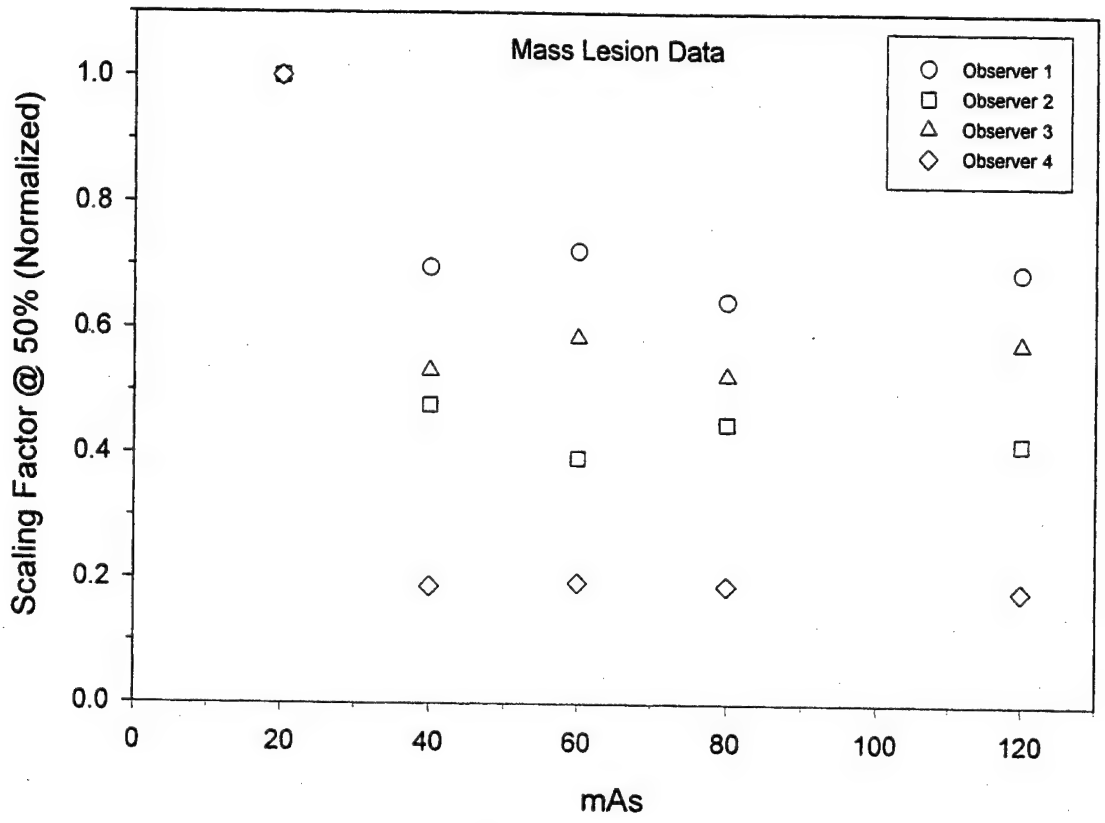


Fig. 5: $SF_{50\%}$ as a function of the mAs for the mass lesion.

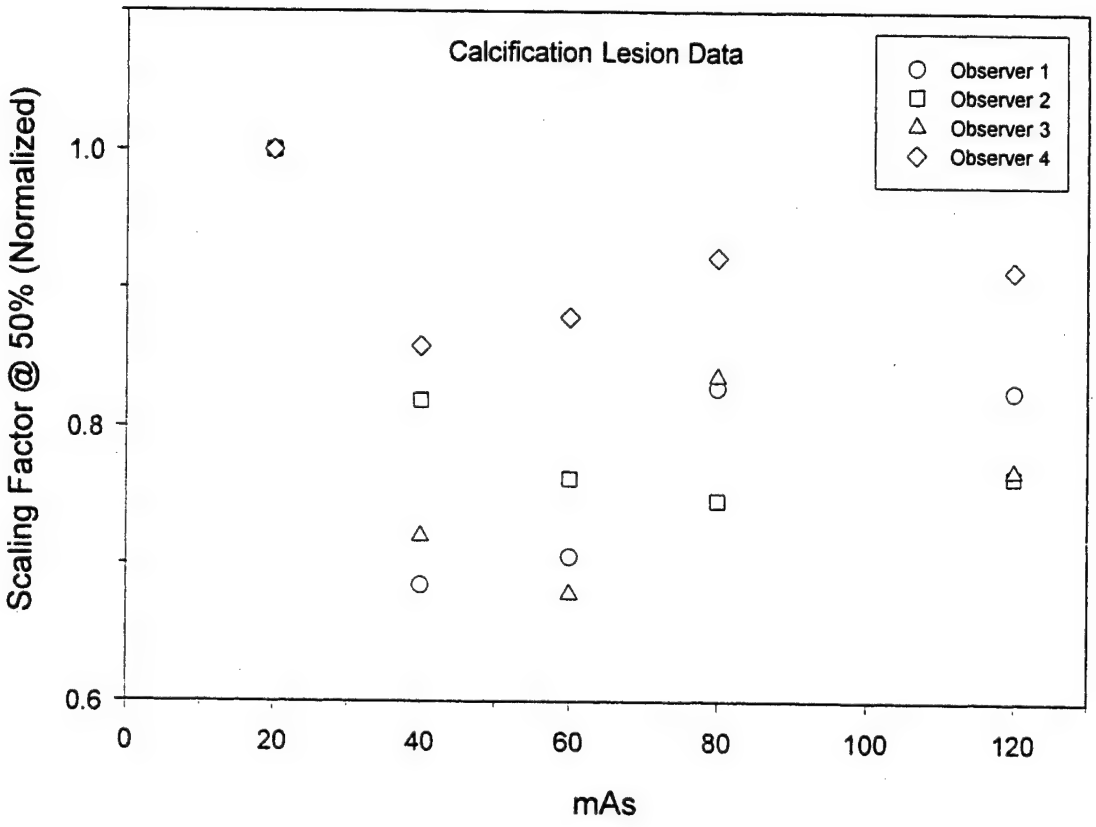


Fig. 6: $SF_{50\%}$ as a function of the mAs for the calcification lesion.

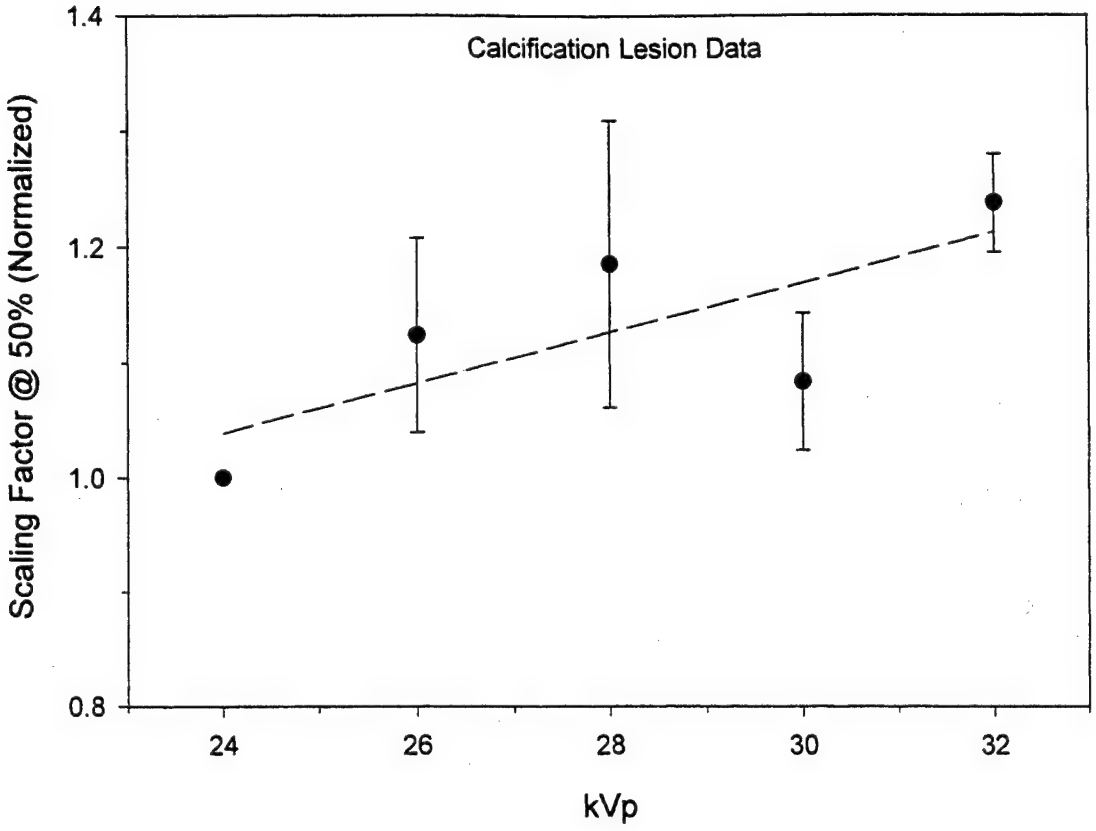


Fig. 7: $SF_{50\%}$ as a function of the kVp for the calcification lesion.

How Good Is the ACR Accreditation Phantom for Assessing Image Quality in Digital Mammography?[1]

Walter Huda, PhD, Anthony M. Sajewicz, MD, Kent M. Ogden, PhD, Ernest M. Scalzetti, MD, David R. Dance, PhD

Rationale and Objectives. The purpose of this study was to evaluate the American College of Radiology (ACR) accreditation phantom for assessing image quality in digital mammography.

Materials and Methods. Digital images were obtained of an ACR accreditation phantom at varying mAs (constant kVp) and varying kVp (constant mAs). The average glandular dose for a breast with 50% glandularity was determined for each technique factor. Images were displayed on a 5 mega-pixel monitor, with the window width and level settings individually optimized for viewing the fibers, specks, and masses in the ACR phantom. Digital images of the ACR phantom were presented in a random manner to eight observers, each of whom indicated the number of objects visible in each image.

Results. Intraobserver variability was greater than interobserver variability for the detection of fibers and specks, but the reverse was true for the detection of masses. As the mAs increased, the number of fibers visible increased from less than one at 5 mAs to all six being visible at 80 mAs. The corresponding number of visible specks increased from 12 to 24, and the number of visible masses increased from 1.25 to about four. Above 26 kVp, object visibility was constant with increasing x-ray tube voltage. Reducing the x-ray tube voltage to 24 kVp, however, reduced the number of visible fibers from six to five, the number of visible specks from 24 to 21.1, and the number of visible masses from four to 3.1. Observer performance was approximately constant for average glandular doses greater than 1.6 mGy, so that the range of lesion detectability in the ACR phantom occurs at doses lower than those normally encountered in clinical practice.

Conclusion. The current design of the ACR phantom is unsatisfactory for assessing image quality in digital mammography.

Key Words. American College of Radiology (ACR) phantom; Digital mammography; Image quality; Observer performance; Radiation dose; Radiographic techniques.

Digital mammography separates the process of image acquisition from any subsequent image display (1). As a result, both image acquisition and display can be optimized independently, which should improve the overall performance of mammography as a diagnostic imaging modality (2,3). Compared with the amount of radiation used in conventional screen-film imaging, the amount of radiation used in digital imaging could be increased or decreased by more than an order of magnitude with no change in the intensity of the displayed image. In addition, the quality of the x-ray beam (ie, half-value layer) used to acquire the digital radiograph may be adjusted by modifying the x-ray tube voltage (ie, kVp) or the choice of the x-ray tube target and filter combination. Any choices made by the operator affect both the image quality and the corresponding radiation dose (4–6).

In screen-film mammography, the American College of Radiology (ACR) accreditation phantom is used to assess the image for quality control purposes (7,8). This phantom contains features such as fibers, speck groups, and

Acad Radiol 2002; 9:764–772

[1] From the Department of Radiology, SUNY Upstate Medical University, 750 E Adams St, Syracuse, NY 13210 (W.H., A.M.S., K.M.O., E.M.S.); and the Department of Medical Physics, Royal Marsden NHS Trust, London, England (D.R.D.). Received January 14, 2002; revision requested February 11; revision received and accepted March 12. Supported in part by U.S. Army grant DAMD 17-00-1003375. **Address correspondence to** W.H.

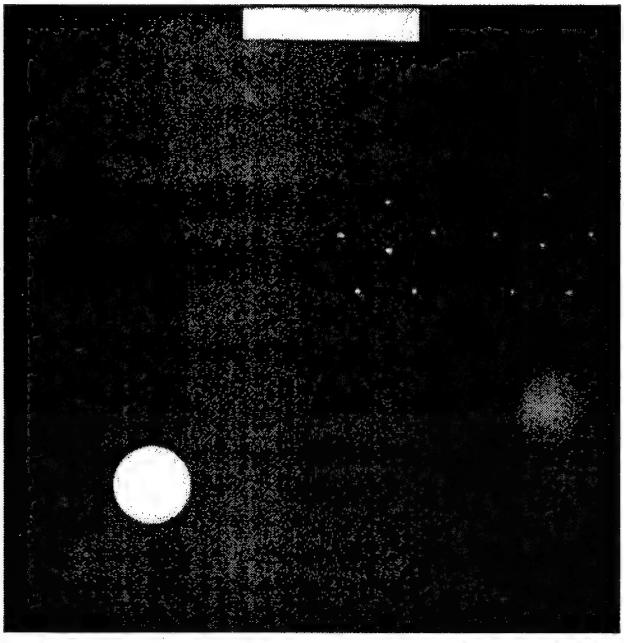


Figure 1. Digital radiograph of the ACR accreditation phantom obtained at 28 kVp and 80 mAs.

masses that simulate lesions of interest during mammography. These lesions are constructed so that their visibility in the resultant radiographic images ranges from the easily visible to the invisible, and, therefore, these lesions straddle the threshold of visibility. Digital mammography requires similar types of phantoms to assess image quality (9–11). An appropriate digital phantom should be capable of indicating changes in image quality over the range of radiographic technique factors expected in clinical practice and, thus, permit a decision regarding whether the digital mammography system can be used for clinical imaging.

In this study, we investigated the detection of simulated lesions (ie, fibers, specks, and masses) in an ACR accreditation phantom as both the x-ray tube output (mAs) and x-ray tube voltage (kVp) were systematically varied. The results of this study shed light on the utility of the ACR accreditation phantom for current digital mammography systems. In addition, the results offer insight as to how imaging performance, in terms of lesion detection in a uniform background, varies with radiographic technique factors. Improved understanding of the effect of radiographic technique factors on lesion visibility will help us to design improved phantoms for digital mammography systems.

MATERIALS AND METHODS

Digital Mammography System

The full-field digital mammography system was a commercial unit manufactured by Lorad (Danbury, Conn). The x-ray spectrum was generated by using a molybdenum target and a molybdenum filter (25 μm) combination. The digital image receptor consisted of a mosaic of twelve 1,600 × 1,600-pixel, charge-coupled devices that were coupled by 2:1 fiberoptic tapers to a large-area, thallium-activated, cesium iodide scintillator plate. The active image area of the image receptor was a field of 18.6 × 24.8 cm, and the image pixel matrix was 4,800 × 6,400. The pixel size at the scintillator surface was 40 μm, resulting in a Nyquist spatial frequency of 12.5 cycles per millimeter. A conventional linear grid (grid ratio, 5:1) was employed to reject scattered x rays.

The hardware platform used to display the images was a standard PC with a diagnostic-quality monitor and video card. The PC was an Optiplex computer (Dell, Round Rock, Tex) operating at 1,000 MHz and with 512 MB of random-access memory with dual 40-GB hard drives. The video system was a 5MP1H display system (Barco, Duluth, Ga) that consisted of a 5-million-pixel gray-scale monitor and a 10-bit (1,024 gray levels) video card; the software package (Osiris, Geneva, Switzerland) was used to display the images. Each pixel had a depth of 14 bits and could cover an intensity range from 0 to 16,383. Digital images had the window width and level settings adjusted to optimize the image display (12). To ensure that images were displayed with the highest possible fidelity, the monitor was calibrated with MediCal software (Barco) and a DTP92Q luminance sensor (X-Rite, Grandville, NJ). This calibration adjusts the geometric accuracy of the displays and calibrates the luminance output of the monitor.

Phantom Exposures

A standard ACR phantom was exposed to obtain digital images at different x-ray tube voltages (kVp) and tube current–exposure time products (mAs). The ACR phantom contains six fibers with diameters of 1.56, 1.12, 0.89, 0.75, 0.54, and 0.40 mm; five speck groups, with six specks in each group, with speck diameters of 0.54, 0.40, 0.32, 0.24, and 0.16 mm; and five masses with decreasing diameters and thickness of 2.00, 1.00, 0.75, 0.50, and 0.25 mm. Figure 1 shows a representative digital radiograph of the accreditation phantom obtained at 28 kVp

and 80 mAs. Manual techniques were used to perform all radiographic exposures.

In one experiment, the x-ray tube voltage was kept constant at 28 kVp, and digital images were generated at tube current–exposure time products of 5–500 mAs. In a second experiment, the tube current–exposure time product was kept constant at 80 mAs, and the x-ray tube voltage was varied in 1-kVp increments between 24 and 34 kVp. In addition, five additional images were obtained at 28 kVp and 80 mAs, which resulted in a total of seven repeat images at these technique factors. These seven repeat images provided information on the experimental reproducibility of each observer, and they permitted the relative sizes of the inter- and intraobserver variabilities to be compared.

The entrance skin exposure and half-value layer were measured according to the recommended protocols of the ACR (7). Entrance skin exposure measurements were converted into corresponding average glandular dose (AGD) values for a compressed breast (thickness, 4.2 cm) with 50% glandularity. At 28 kVp and 80 mAs, the AGD was 2.16 mGy and directly proportional to the mAs. Figure 2 shows how the AGD varies with the x-ray tube voltage at a constant 80 mAs. The dotted line in Figure 2 is a least-squares fit to a second-order polynomial, and the equation permits the AGD to be determined for any selected x-ray tube voltage.

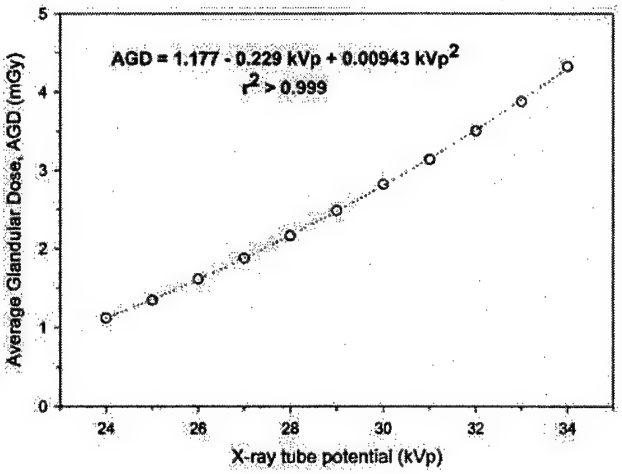


Figure 2. AGD as a function of x-ray tube potential obtained at a constant 80 mAs for an average-size breast with 50% glandularity.

Table 1
Scoring Summary Used to Assess Feature Visibility

Feature	Partial Scores	Maximum Score
Fibers	0.25, 0.5, 0.75	6
Specks	None	30
Masses	0.25, 0.5, 0.75	5

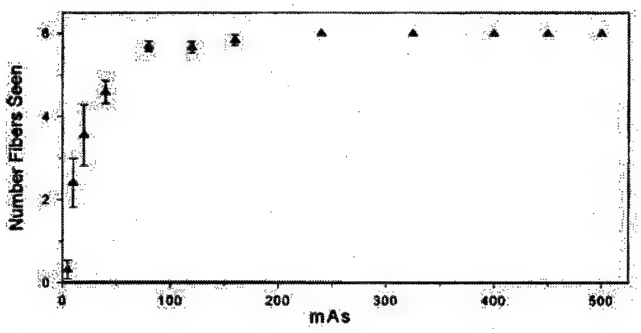
Observer Study

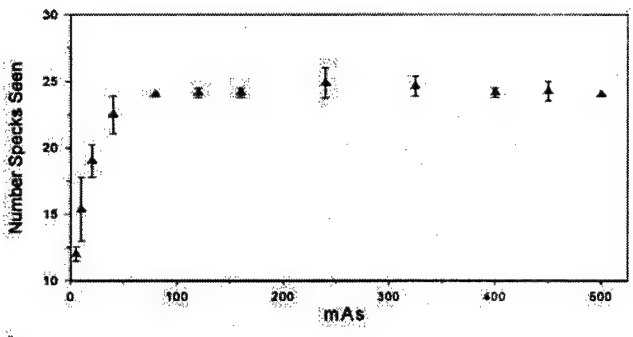
Each image was individually optimized for viewing the fibers, microcalcification specks, and masses. The average pixel intensities of the mid-size fibers (1.12 and 0.89 mm) and masses (1.0 and 0.75 mm) were recorded together with the adjacent background regions. These values were used to manually set the approximate window width and level display settings in the imaging software. An experienced operator then "fine-adjusted" the window width and level settings to optimize further the display of each ACR accreditation phantom image.

Eight readers were used, including two experienced medical physicists, two radiologists, two radiology residents, and two technologists. All non–medical physicist readers were given a short orientation course on the ACR accreditation phantom; this course was developed based on the material provided in the ACR mammography handbook. Images of the ACR phantom were examined and scored according to what was actually seen of the three types of features. Observers were not allowed to make additional adjustments to these displayed images. Table 1 summarizes the scoring scheme; partial scores were allowed when a lesion was not wholly visualized. We did not consider any artifacts that might have been visible on the phantom images. The scoring scheme used in this study differed from that recommended by the ACR, because our objective was to study how lesion visibility varied with the x-ray techniques that were used.

A total of 28 phantom images were obtained, consisting of 11 in the kVp series, 12 in the mAs series, and five repeat examinations. These 28 images were presented to the observers randomly and in a single setting. The average time required by the eight observers to read these images was 29.6 minutes ± 2.5 (minimum, 25 minutes; maximum, 33 minutes).

Results are presented as either absolute or relative values. In the absolute format, the number of lesions in each category reported as being visible by each observer was recorded, and the average (± standard deviation) was then computed. In the relative format, the number of lesions in each category reported as being visible was sub-





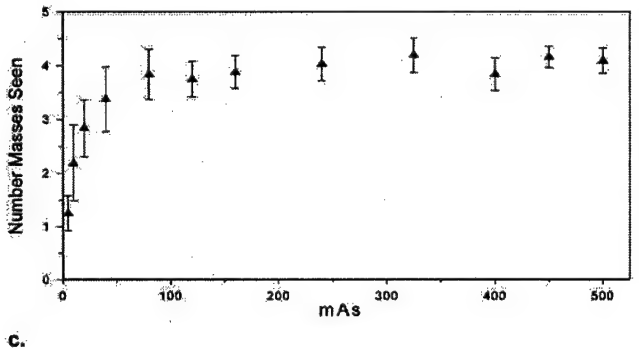


Figure 3. Absolute number of **(a)** fibers, **(b)** specks, and **(c)** masses seen in the ACR phantom image versus mAs. Each datum is the mean value for eight observers, and error bars are the corresponding standard deviations.

Table 2
Precision Summary for the Eight Observers Reading Seven Repeat Images

Observer*	Fibers	Specks	Masses
1 (Rad)	5.43 ± 0.40	23.86 ± 0.38	3.50 ± 0.29
2 (Phy)	5.79 ± 0.27	23.86 ± 0.38	3.75 ± 0.38
3 (Phy)	5.61 ± 0.38	23.86 ± 0.38	3.86 ± 0.13
4 (Tech)	5.61 ± 0.35	23.86 ± 0.38	3.93 ± 0.12
5 (Tech)	5.57 ± 0.28	23.86 ± 0.38	3.79 ± 0.17
6 (Rad)	5.64 ± 0.32	23.71 ± 0.49	2.79 ± 0.34
7 (Res)	5.75 ± 0.14	24.00 ± 0.58	4.18 ± 0.19
8 (Res)	5.68 ± 0.35	24.71 ± 1.38	3.32 ± 0.57
Average observer score	5.64 ± 0.11	23.97 ± 0.31	3.64 ± 0.43
Average intraobserver variability	0.31	0.54	0.27

Note.—Images were obtained at 28 kVp and 80 mAs. Data presented as mean ± standard deviation.

*Phy = medical physicist, Rad = radiologist, Res = resident, Tech = technologist.

tracted from the corresponding average number of lesions seen by that observer at the 28 kVp and 80 mAs setting. Use of the relative format enabled interobserver differences to be reduced.

RESULTS

Experimental Precision

The precision of the window width and level scheme was evaluated with the seven repeat experiments at 28 kVp and 80 mAs. The average window level for these images was 2,402 ± 24; the average window width setting was 436 ± 106. These data indicate that the relative precision for setting the level was approximately 1% of the average pixel intensity. For the window width, however, the relative precision was approximately 24% of the average window width.

Table 2 summarizes the experimental data for all eight observers in reading seven repeat images obtained at 28 kVp and 80 mAs. The standard deviations for each category include inter- and intraobserver variabilities. Reproducibility of the contrast-to-noise ratio (CNR) in these images was approximately 3% (13) and was considered to be negligible compared with the variability in the observers' performances.

For fibers, the interobserver variability (± 0.11 fibers) was lower than the average observer precision (± 0.31 fibers) by a factor of approximately three. For specks, the interobserver variability (± 0.31 specks) was smaller than the average observer precision (± 0.54 specks). For masses, however, the interobserver variability (± 0.43 masses) was much larger than the average observer precision (± 0.27 masses). Thus, the relative magnitude of inter- and intraobserver variabilities depended on the detection task.

Observer Performance and Technique Factors

Figure 3 shows the absolute number of objects detected by the eight observers as a function of the mAs

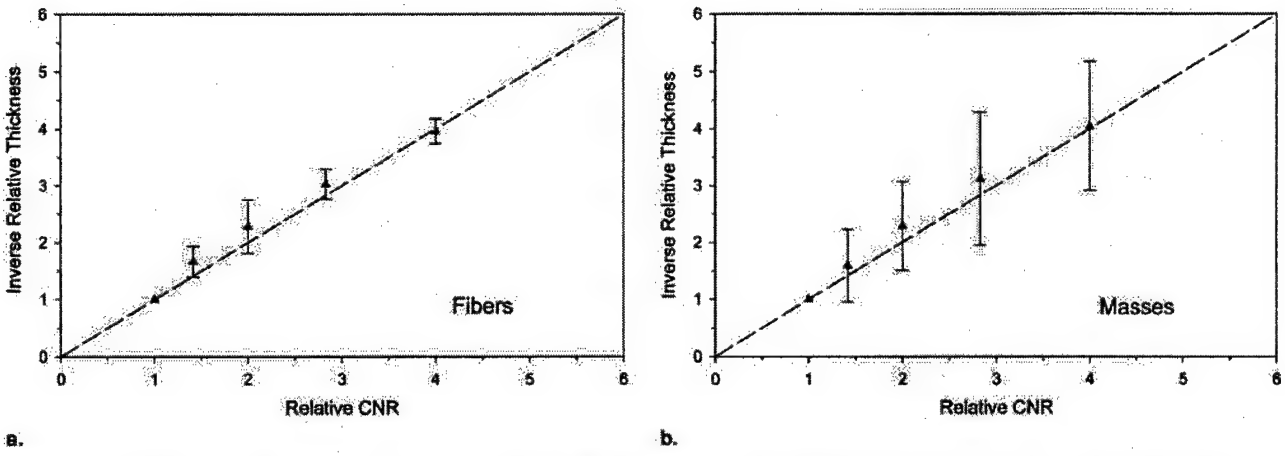


Figure 4. Inverse relative thickness of **(a)** fibers and **(b)** masses at the threshold of visibility plotted as a function of the image CNR computed from the change in mAs used to acquire the phantom images. The dashed line is the line of equality, where the inverse relative thickness is equal to the relative CNR.

value. Figure 3a shows the same results for the fibers; a rapid increase, from an average of 0.31 fibers visible at 5 mAs to more than 5.5 fibers visible at 80 mAs, can be seen. At 240 mAs and above, all observers could see all six fibers. Figure 3b shows that the number of visible specks increased from an average of 12 at 5 mAs to a "plateau value" of 24.0 at 80 mAs. Figure 3c shows that the number of visible masses increased from 1.25 at 5 mAs to a plateau value of approximately four visible masses at 80 mAs.

It is interesting to compare how the size of each lesion at the threshold of detection changed as the x-ray tube output increased from 5 to 80 mAs. The ACR phantom manufacturer provides information regarding the size of each type of lesion, and by plotting size as a function of lesion number, we could convert the lesion visibility (depicted as the ordinate in Fig 3) to the corresponding lesion size. The relative CNR ($CNR_{relative}$) was calculated from the change in the mAs values used to acquire each phantom image; because this imaging system is approximately quantum noise limited (13), doubling the mAs increases the $CNR_{relative}$ by 41%. Figure 4 shows how the inverse of relative fiber and mass thickness varies as a function of the $CNR_{relative}$; all the data have been normalized to the lesion threshold thickness value obtained at 5 mAs. Figure 5 shows how the inverse of relative speck cross-sectional area varies with $CNR_{relative}$.

Figure 6 shows the absolute number of objects detected by the eight observers as a function of x-ray tube voltage. These data show that for each type of object, detection was essentially constant between 26 and 34

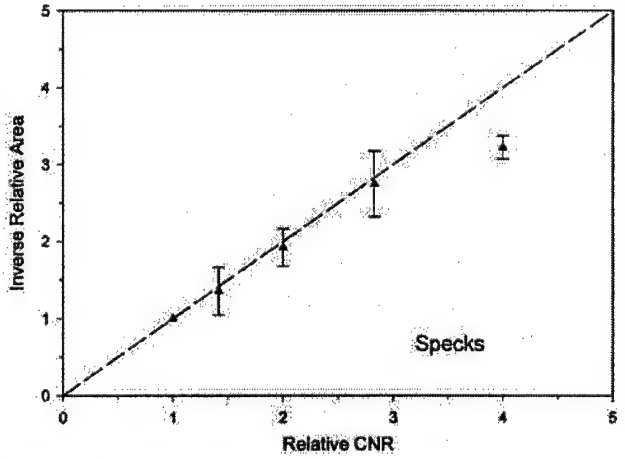


Figure 5. Inverse relative cross-sectional area of microcalci fication specks at the threshold of visibility plotted as a function of the image CNR computed from the change in mAs used to acquire the phantom images. The dashed line is the line of equality, where the inverse relative cross-sectional area is equal to the relative CNR.

kVp, reaching a plateau level that corresponded to the best performance obtained at the highest mAs value shown in Figure 3. Reducing the x-ray tube voltage to 24 kVp, however, reduced the number of visible fibers from six to five, the number of visible specks from 24 to 21.1, and the number of visible masses from 4 to 3.1.

Observer Performance and AGD

Figure 7 shows the relative performance of the observers as a function of the AGD when results from

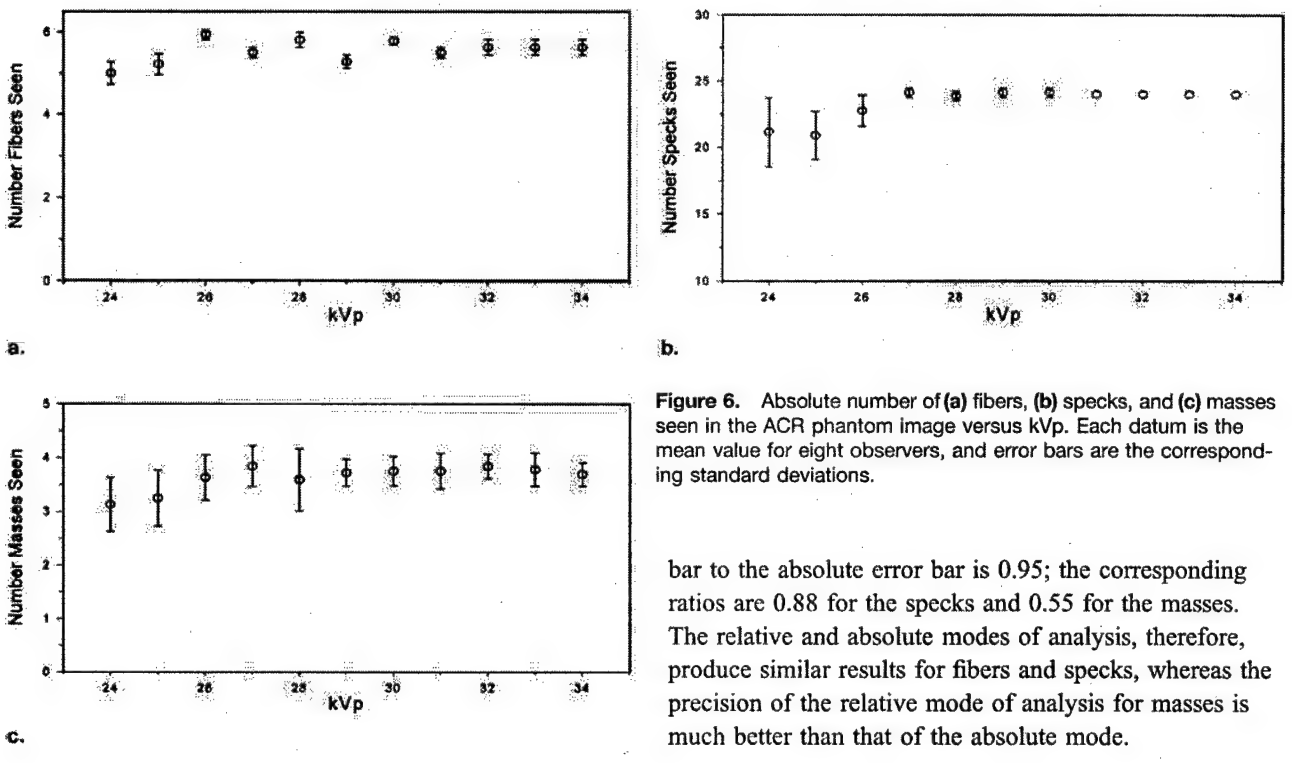


Figure 6. Absolute number of **(a)** fibers, **(b)** specks, and **(c)** masses seen in the ACR phantom image versus kVp. Each datum is the mean value for eight observers, and error bars are the corresponding standard deviations.

both mAs and x-ray tube voltage experiments are included. At low AGDs, a close correlation between dose and image quality is observed whether or not the dose is modified by changing the mAs or the x-ray tube voltage. For each type of lesion, an estimate was made of the AGD at which 90% of the maximum numbers of lesions were actually visualized. Figure 7 shows that this level occurred at 1.1 mGy for fibers, 1.4 mGy for masses, and 1.6 mGy for specks. Lesion visibility thus varies in the low-dose region with AGDs of less than 1.6 mGy. Little variation was found in observer performance at doses that are expected in clinical mammography (~1.6 mGy) (14).

Figure 7 also plots the relative number of objects seen by each observer and, therefore, minimizes the effect of interobserver variability. To illustrate how this "relative" mode of presentation affects the observed errors, we examine the data at 28 kVp and 20 mAs. At this technique factor, an average of 3.6 fibers, 19.0 specks, and 2.8 masses were detected; because the rate of change in detection performance was high, the corresponding errors should be maximized. Figure 3 shows the absolute values for the three features at a technique of 20 mAs, and Figure 7 shows the corresponding relative values at a dose of 0.54 mGy. For the fibers, the ratio of the relative error bar to the absolute error bar is 0.95; the corresponding ratios are 0.88 for the specks and 0.55 for the masses. The relative and absolute modes of analysis, therefore, produce similar results for fibers and specks, whereas the precision of the relative mode of analysis for masses is much better than that of the absolute mode.

DISCUSSION

In this study, the display was individually optimized for each acquired image when viewing the lesions in the ACR phantom. The precision experiments showed that the method employed to optimize the image display had excellent reproducibility for setting the window level value (1%) but much greater variability for setting the window width (24%). The average window width for displaying the images generated at 28 kVp and 80 mAs was only 436 pixel intensity values, which is only 2.7% of the total dynamic range of the digital detector. Because the ACR phantom is relatively uniform in composition and contains only low-contrast objects, the x-ray pattern has a small dynamic range. Therefore, that considerable variability was found in the window width setting used to display images acquired at the same technique factor is not surprising, and it indicates that the window width is not critical for visualizing these structures.

The detector uses a 14-bit system to meet the stringent dynamic range requirements of digital mammography (15). As a result, the maximum signal intensity corresponds to a maximum pixel intensity of 16,383, and changing the tube current–time product from 5 to 500 mAs makes use of this wide dynamic range. By compari-

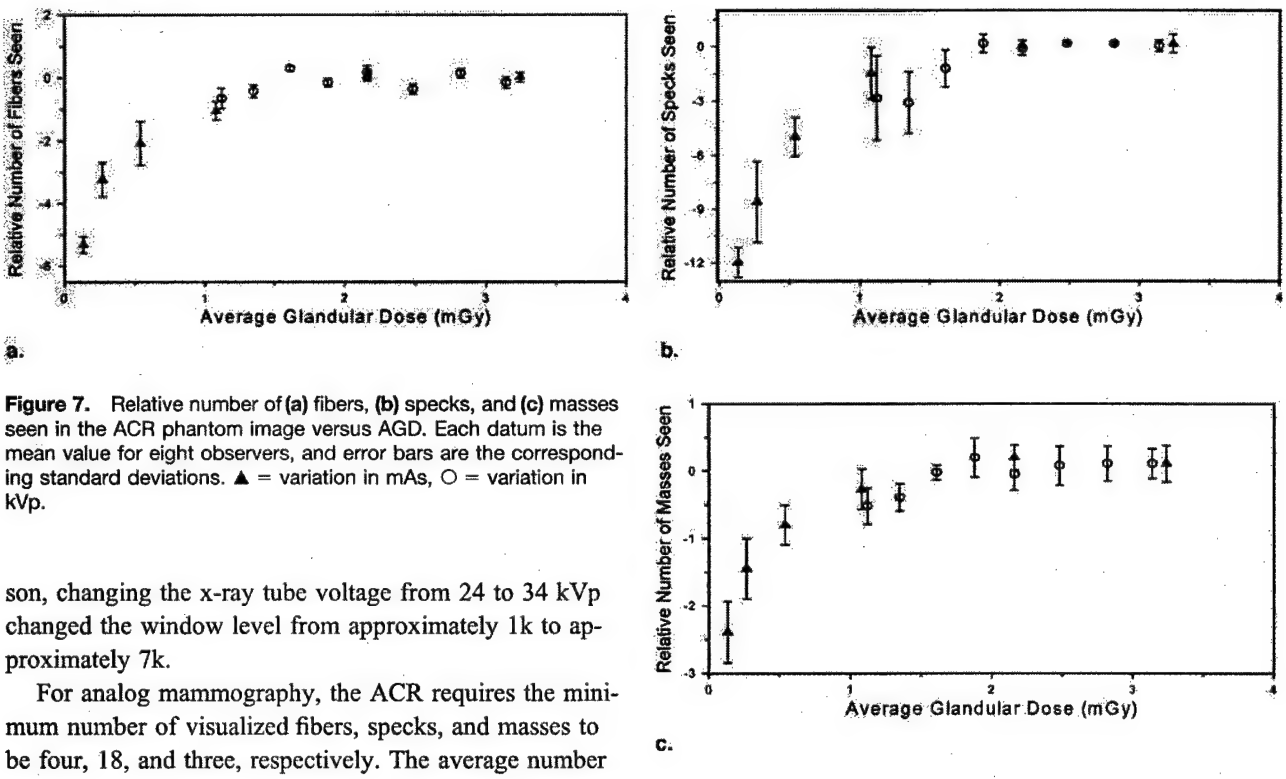


Figure 7. Relative number of **(a)** fibers, **(b)** specks, and **(c)** masses seen in the ACR phantom image versus AGD. Each datum is the mean value for eight observers, and error bars are the corresponding standard deviations. ▲ = variation in mAs, ○ = variation in kVp.

son, changing the x-ray tube voltage from 24 to 34 kVp changed the window level from approximately 1k to approximately 7k.

For analog mammography, the ACR requires the minimum number of visualized fibers, specks, and masses to be four, 18, and three, respectively. The average number of objects visible in a digital image of the ACR phantom acquired at 28 kVp and 80 mAs was 5.6 for fibers, 24.0 for specks, and 3.6 for masses. This performance level is higher than that expected for screen-film mammography, in which images of the ACR phantom typically show, approximately, five fibers, 21 specks, and three and a half masses. Comparing the present results with conventional phantom scores is problematic, because the latter generally include penalties allocated for the presence of image artifacts. Most artifacts on the ACR phantom, however, are specks and rarely include masses or fibers. The performance of this digital system is slightly superior to that of screen-film mammography for the detection of low-contrast objects such as fibers and masses, which is in agreement with the findings of a recent study by Undrill et al (16). Our results also suggest that detection of microcalcification specks with digital technology is slightly superior to that with screen-film mammography, which is in agreement with the findings of a recent study by Cowen et al (17). Digital mammography is expected to be superior to screen-film mammography because of its ability to independently optimize image acquisition and to maintain excellent image contrast by digital manipulation of the display (18).

The choice of x-ray tube voltage affects both dose and image quality and depends on the thickness and composition of the compressed breast (19). In addition, digital mammography likely may require different spectra from those normally used in screen-film mammography (20). At constant mAs, reducing the x-ray tube voltage generally increases the subject contrast and image noise. The reduced detection performance observed at 24 and 25 kVp suggests that increased noise is more important than the corresponding improvement in subject contrast for all three types of lesions in the uniform background of the ACR phantom.

Variability in detection performance when observing lesions of the type considered in this study are normally classified as interobserver, intraobserver, and sample variance (21,22). Previous investigations have considered these sources of error for classical contrast-detail experiments. Their results have generally shown that interobserver variability is the most important, but these results depend on the size of the lesion being evaluated and on the type of reader performing the study. Fibers and specks are well-defined objects and permit different observers to agree on the number visible. Masses, however, are more subjective, and different observers employ different criteria (ie, thresholds) for determining their degree of visibility. In the future, phantom image quality likely

will be assessed with computerized, objective measures that can be expected to eliminate inter- and intraobserver variability (23).

Figure 7 illustrates the direct link between dose and image quality in digital mammography (24). Observer performance improved with increasing radiation dose up to an AGD of 1.6 mGy. The AGD in screen-film mammography is at least 1.5 mGy, and current full-field digital mammography systems operate with techniques similar to those of screen-film systems. Present regulations require the AGD for a normal-size breast (50% fibroglandular tissue and 50% adipose tissue) to be no more than 3 mGy. Consequently, expecting a phantom in digital mammography to have a range of detectability over the dose range of approximately 1½ to 3 mGy is reasonable. The data shown in Figure 7 indicate that for the ACR phantom, observer performance was constant between 1.6 and 3.0 mGy. Current digital mammography systems are expected to be quantum noise limited and would be expected to produce similar results. Our findings therefore imply that the ACR phantom is unsatisfactory for assessment of image quality in digital mammography.

Given that digital mammography is expected to be superior to screen-film mammography, it is unsurprising that the design of the ACR phantom requires modification (25). The data shown in Figures 3 and 6 indicate that the fifth mass and the fifth group of specks were essentially invisible, presumably because the lesion signal-to-noise ratio was low. Detection for all three types of lesions (fibers, specks, and masses) showed similar characteristics between 5 and 80 mAs. Data shown in Figure 4 indicate that for fibers and masses, $CNR_{relative}$ is proportional to the inverse of the lesion thickness, whereas the data shown in Figure 5 demonstrate that for specks, the $CNR_{relative}$ is proportional to the inverse of the speck cross-sectional area. The findings reported in this study could be used to guide the construction of a quality-control phantom that would have the appropriate range and sensitivity for current types of digital mammography imaging systems.

ACKNOWLEDGMENTS

The authors are grateful to Zhenxue Jing, PhD, for assistance with the experimental work and useful discussions on digital mammography. Lorad provided access to the digital mammography imaging system and the ACR accreditation phantom. The authors also thank Marsha L. Roskopf, BS, Elizabeth Bertrand, RT, Robert Dixon, MD, Dennis Hseuh, MD, and Kristen Lieberman, MD, for reading the images.

REFERENCES

1. Yaffe M. Digital mammography. In: Haus AG, Yaffe MJ, eds. Syllabus: a categorical course in diagnostic radiology physics: physical aspects of breast imaging—current and future considerations. Oak Brook, Ill: Radiological Society of North America, 1999; 229–238.
2. Liu H, Fajardo LL, Barrett JR, Baxter RA. Contrast-detail detectability analysis: comparison of a digital spot mammography system and an analog screen-film mammography system. Acad Radiol 1997; 4:197–203.
3. Pisano E. Current clinical status of full-field digital mammography. In: Haus AG, Yaffe MJ, eds. Syllabus: a categorical course in diagnostic radiology physics: physical aspects of breast imaging—current and future considerations. Oak Brook, Ill: Radiological Society of North America, 1999; 239–241.
4. Gingold EL, Wu X, Barnes GT. Contrast and dose with Mo-Mo, Mo-Rh, and Rh-Rh target-filter combinations in mammography. Radiology 1995; 195:639–644.
5. Young KC, Ramsdale ML, Rust A. Dose and image quality in mammography with an automatic beam quality system. Br J Radiol 1996; 69:555–562.
6. Thilander-Klang AC, Ackerholm PHR, Berlin IC, et al. Influence of anode-filter combinations on image quality and radiation dose in 965 women undergoing mammography. Radiology 1997; 203:348–354.
7. American College of Radiology. Mammography quality control manual. Reston, Va: American College of Radiology, 1999.
8. McLean D, Eckert M, Heard R, Chan W. Review of the first 50 cases completed by the RACR mammography QA programme: phantom image quality, processor control and dose considerations. Australas Radiol 1997; 41:387–391.
9. Roehrig H, Yu T, Krupinski E. Image quality control for digital mammographic systems: initial experience and outlook. J Digit Imaging 1995; 8:52–66.
10. Kimme-Smith C, Lewis C, Beifuss M, Williams MB, Bassett LW. Establishing minimum performance standards, calibration intervals, and optimal exposure values for a whole breast digital imaging mammography unit. Med Phys 1998; 25:2410–2416.
11. Chakrabarti K, Kaczmarek R, Thomas J, Brinkwine B, Loscocco M. Effects of CNR and SNR on phantom image quality in GE Senograph 2000D full field digital mammography system. Med Phys 2001; 28:1247.
12. Chakrabarti K, Thomas JA, Kaczmarek RV, Waynant RW, Loscocco MF. Optimization of viewing conditions and phantom image quality evaluations on GE DMR and full-field digital mammography systems. J Digit Imaging 2000; 13:226–227.
13. Huda W, Sajewicz AM, Ogden KM, Dance DR. Experimental investigation of the dose and image quality characteristics of a digital mammography imaging system. Med Phys (submitted).
14. Lavoy T, Huda W, Ogden K. Radiographic techniques in screen-film mammography. J Appl Clin Med Phys (in press).
15. Maidment ADA, Fahrig R, Yaffe MJ. Dynamic range requirements in digital mammography. Med Phys 1993; 20:1621–1633.
16. Undrill PE, O'Kane AD, Gilber FJ. A comparison of digital and screen-film mammography using quality control phantoms. Clin Radiol 2000; 55:782–790.
17. Cowen AR, Launders JH, Jadav M, Brettle DS. Visibility of microcalcifications in computed and screen-film mammography. Phys Med Biol 1997; 42:1533–1548.

18. Haus AG, Yaffe MJ. Screen-film and digital mammography: image quality and radiation dose considerations. Radiol Clin North Am 2000; 38:871–898.
19. McParland BJ, Boyd MM. A comparison of fixed and variable kVp technique protocols for screen-film mammography. Br J Radiol 2000; 73:613–626.
20. Kimme-Smith C. New digital mammography systems may require different x-ray spectra and, therefore, more general normalized glandular dose values. Radiology 1999; 213:7–10.
21. Loo LN, Doi K, Ishida M, et al. An empirical investigation of variability in contrast-detail diagram measurements. SPIE 1983; 419:68–76.
22. Cohen G, McDaniel DL, Wagner LK. Analysis of variations in contrast-detail experiments. Med Phys 1984; 11:469–473.
23. Castellano-Smith AD, Castellano-Smith IA, Dance DR. Objective assessment of phantom image quality in mammography: a feasibility study. Br J Radiol 1998; 71:48–58.
24. Haus AG, Yaffe MJ, Hendrick ER, Butler PA, Wilcox PA, Bansal S. Relationship between phantom failure rates and radiation dose in mammography accreditation. Med Phys 2002; 28:2297–2301.
25. Bijkerk KR, Thihssen MAO, Arnoldussen TJM. Modification of the CDMAM contrast-detail phantom for image quality evaluation of full field digital mammography systems. In: Yaffe MJ, ed. Proceedings of the 5th international workshop on digital mammography. Madison, Wis: Medical Physics, 2001; 633–640.

ERRATUM

"Prediction of Treatment Response of Head and Neck Cancers with P-31 MR Spectroscopy from Pretreatment Relative Phosphomonoester Levels." Acad Radiol 2002; 9:688–694

The authors also gratefully acknowledge support from National Institutes of Health grant P01 CA05826-038A1.

Comparison of experimental and theoretical assessments of detail visibility in digital mammography

David R Dance[1], Roger A Hunt[1], Anthony M Sajewicz[2], Walter Huda[2], Kent M Ogden[2], Michael Sandborg[3] and Gudrun Alm Carlsson[3]

[1]Joint Department of Physics, The Royal Marsden NHS Trust, Fulham Road, London SW3 6JJ, United Kingdom
[2]Department of Radiology, SUNY Upstate Medical University, 750 E Adams Street, Syracuse, NY 13210, USA
[3]Department of Radiation Physics, IMV, Faculty of Health Science, Linköping University, SE-581 85 LINKÖPING, Sweden
d.dance@icr.ac.uk

Abstract. An experimental and modelling study is being made of the influence of tube voltage, target material and exposure on the performance of digital mammography systems. Digital images of the ACR accreditation phantom at 80 mAs, 25-32 kV and at 28 kV, 5-500 mAs were read by eight observers, and the numbers of fibres, specks and masses visible determined. The computer model simulates photon transport through phantom, anti-scatter grid and image receptor. It calculates image pixels and the signal-to-noise ratio per pixel (SNR) for the phantom details. For exposures below 100 mAs, the numbers of fibres and masses visualised were found to be consistent with a constant SNR threshold for detection. For the visualisation of specks, the product of SNR and speck area was approximately constant. At higher mAs, the number of objects visualised was little influenced by exposure, due to the limited dynamic range of the phantom. The results validate the use of computational models to predict performance for simple detection tasks against a uniform background.

1. Introduction

In digital mammography it is possible to modify the image prior to display so that contrast requirements associated with screen-film mammography may be relaxed. In addition, because of the improved dynamic range and DQE, it may be possible to reduce the dose for imaging with a given radiation quality or to use a different radiation quality, with further dose saving.

We are making an experimental and computer modeling study of the influence of user-controllable parameters such as exposure, tube voltage and target material on the performance of digital mammography systems. In this paper[1] we present the results of measurements and calculations of detail visibility for the ACR accreditation phantom (Gammex-RMI, Middleton, WI) imaged with a LoRad (Danbury, CT) digital mammography system at a series of tube voltage and exposure values.

[1] This work is supported in part by US Army grant No DAMD 17-00-1003375.

2. Materials and Methods

An ACR accreditation phantom was exposed using the LoRad digital mammography system and an x-ray spectrum from a Mo target/Mo filter. Images were acquired at 28 kV for 12 values of the exposure between 5-500 mAs and at 80 mAs for 8 values of the tube voltage between 25-32kV. The phantom contains sets of details which simulate fibrils (six nylon filaments with diameters between 0.40 - 1.56 mm), calcifications (six groups of alumina specks with sizes between 0.16 - 0.54 mm) and masses (water density objects of thicknesses between 0.25 - 2.00 mm). The soft-copy images were viewed under controlled conditions by eight readers and the number of objects of each type which were visible were scored [1].

The model used was based on a Monte Carlo computer program developed for modeling mammographic systems [2], and extended to treat the 'patient' as an array of voxels. For this case each voxel was composed of polymethyl methacrylate apart from a 5 mm layer of wax which contained the test details. The phantom was 45 mm thick and rested on a 2 mm thick carbon fibre support. The anti-scatter grid had ratio 5 and 31 lines.cm^{-1}. The x-ray spectra were adapted so that they matched measured HVL values. The image receptor was 73 mg.cm^{-2} CsI. The program could calculate mean glandular breast dose, entrance air kerma, the energy imparted per image pixel, image contrast and signal-to-noise ratio per pixel (SNR). In this way the SNR per pixel was calculated for each detail size and type, exposure and tube voltage.

In its present form, the model makes no allowance for image unsharpness. This will affect the estimation of contrast for the smallest details and of the fractional noise per pixel, both of which will decrease.

3. Results and Discussion

Figure 1 shows the calculated (normalised) SNR per pixel for the smallest detail detected by the observers at 28 kV for exposures between 5 and 160 mAs. The results are for the detection of nylon filaments (left) and masses (right). The central value of the SNR at each mAs corresponds to the size of the smallest detectable detail averaged over observers. The error bars represent one standard deviation. The different data points thus correspond to the detection of objects of different sizes. The size of the smallest detectable object and the associated image contrast become smaller as the exposure increases because there is an associated decrease in the image noise.

In both cases, for exposures of 80 mAs and below, the SNR for the smallest detail visualised is consistent with a constant value. In other words, objects are detected when the SNR exceeds a threshold value. Above 80 mAs, some observers see all 6 filaments and the calculated SNR increases and no longer represents the threshold value. A similar comment applies to the masses. In this case, only 4 out of 5 of the masses were generally seen at the higher exposures, which may relate to the fixed display window used for the experiments and the low contrast of the smallest mass (1% at 28 kV). The absolute values of the threshold SNR per pixel are different for the masses and filaments. This is because the imaging tasks are different. The first

involves the detection of long thin cylinders of low contrast (where the detection task may involve integration along the length of the cylinder) whereas the second involves the detection of circular, medium to large diameter low contrast objects.

The results for the specks differed from those for the filaments and masses. In this case, the objects are small and circular and it was found that for low mAs values, the product of the SNR per pixel and the area for the smallest detectable speck was approximately constant. In other words for small round objects, the total signal (integrated over the image of the detail) is strongly correlated with dectability.

The calculations of SNR for the smallest details detected in the experiments where the kV was varied show similar behaviour to those where the mAs was varied.

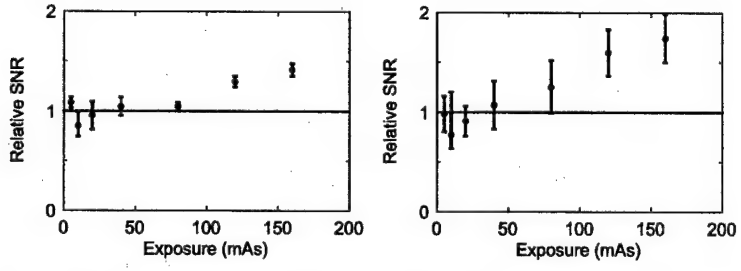


Fig. 1. Relative values of the calculated SNR corresponding to the smallest detectable filament and mass (*left and right figures*) as a function of exposure at 28 kV. Values are normalised to the mean of the first five data points (*horizontal line*).

4. Conclusion

The SNR calculated by the computer model can be used to predict the detectability of simulated fibrils, calcifications and masses against a uniform background. Since the model can also estimate mean glandular dose, it forms a useful tool for the optimisation of the exposure conditions (mAs, kV, target material and filter) so that dose can be minimised for these simple detection tasks. How these tasks relate to the detection of abnormality against a structured background needs to be established.

References

1. Huda, W., Sajewicz, A.M., Ogden K.M., Scalzetti, E.M., Dance, D.R.: How good is the ACR accreditation phantom for assessing image quality in digital mammography? Acad. Radiol. (2002) In press
2. Dance, D. R., Thilander-Klang, A., Sandborg, M., Skinner, C. L., Castellano Smith, I., and Alm Carlsson, G.: Influence of anode/filter material and tube potential on contrast, signal-to-noise ratio and average absorbed dose in mammography: a Monte Carlo study, Br. J. Radiology 73 (2000) 1056-1067

Comparison of objective and subjective methods to assess imaging performance in digital mammography

Walter Huda[1], Kent Ogden[1], Anthony Sajewicz[1], David Dance[2], Ernest Scalzetti[1]

[1]Radiology, SUNY Upstate Medical University, Syracuse NY 13210
[2]Joint Department of Physics, The Royal Marsden NHS Trust, London SW3 6JJ
hudaw@mail.upstate.edu

Abstract. Digital images were acquired of an anthropomorphic breast phantom (Rachel) at mAs values that ranged from 40 to 120 mAs. Digital mammograms were obtained with/without added lesions, which permitted the generation of a digital version of the lesion alone. The lesion was added back to the digital image of the Rachel phantom at varying levels of contrast. A subjective method of lesion visibility was compared to an objective method used Four Alternate Forced Choice (4AFC) methodology. Both methods showned observer performance for detecting a mass lesion to be *independent* of x-ray tube output between 40 and 120 mAs. The objective approach results in a higher precision, but alo requires many more images.

1. Introduction

The assessment of imaging performance in digital mammography is important for selecting the optimum technique factors (kVp/mAs) and for evaluating the utility of new image processing algorithms. For example, it is of considerable interest to determine the effect of x-ray tube output (mAs) on observer performance [1]. In this study, we compared a subjective imaging performance approach with an objective method for assessing how lesion detection changes with increasing mAs in digital mammography.

2. Method

2.1 Digital image acquisition

Digital images were obtained of an anthropomorphic breast phantom (Rachel) using a Lorad Full Field Digital Mammography imaging system at four mAs values (i.e., 40, 60, 90 and 120 mAs). Digital mammograms were made with/without added lesions, which permitted the generation of a digital version of the lesion alone. The 1 cm diameter lesion was added back to the digital image of the Rachel phantom alone at varying levels of contrast. The threshold contrast level for detection was investigated as a function of the mAs using three observers.

2.2 Observer assessment

For a given mAs, the lesion intensity (Scaling Factor/SF) was modified to produce six images where lesion visibility (i.e., contrast) ranged from the "extremely difficult" to "easily seen". Five copies of these 24 different images (4 mAs values & 6 SF factors) were generated to produce a series of 120 images which were presented to each observer in a single reading. Observers specified a probability of a lesion being present on a scale ranging from 0 to 100% which permitted the subjective scaling factor value for a 50% probability to be obtained at each mAs value [2].

The objective method used Four Alternate Forced Choice (4AFC) methodology with an observer identifying which one of four images actually contained the lesion. The results permitted the objective scaling factor to be determined as the value at which the observer accuracy was 92% at each mAs value. An accuracy of 92% corresponds to a theoretical signal to noise ratio of 2.5 (d') [3].

3. Results

3.1 Objective results

Figure 1 (left) shows the results obtained in the objective 4 AFC experiment, where SF for 92% accuracy is plotted as a function of mAs, with each datum the average (± standard deviation) for three observers. The coefficient of determination for the least squares linear fit to the four data points (r^2) was 0.59.

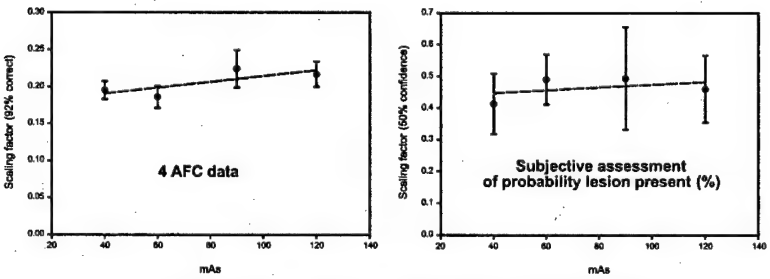


Figure 1. Comparison of objective (4 AFC) and subjective assessments of observer performance.

These data also show no significant correlation between SF(92%) and mAs. The average SF for a 92% correct score was 0.20, and the average coefficient of variation at the four mAs values was 8%.

3.2 Subjective results

Figure 1 (right) shows the results obtained in the subjective experiment, where SF at 50% observer confidence level is plotted as a function of mAs. The coefficient of determination for the least squares linear fit to the four data points (r^2) was 0.18, and these data clearly show that changing the mAs had no effect on observer performance. The average SF for a confidence score of 50% was 0.46 for this mass lesion, and the average coefficient of variation at the four mAs values was 24%.

4. Discussion

The results of both subjective and objective methods of assessing observer performance gave similar trends of performance as a function of the mAs used to acquire these digital images. There was no evidence of any significant change in observer performance as the x-ray tube output increased by a factor of three from 40 to 120 mAs.

The precision (coefficient of variation) in the subjective method was a factor of three larger than the corresponding precision with the 4 AFC approach. In part, this is due to the fact that the subjective method required an observer to read 30 images, whereas the 4 AFC approach required 128. The absolute level of scaling factor differed markedly between the two approaches; the subjective method requiring an SF value of 0.46 to achieve a confidence score of 50%, whereas a SF of 0.21 resulted in a observers achieving a 92% correct score.

Both subjective and objective modes of evaluating observer performance have advantages. The 4AFC method is clearly superior, but is also much more labor intensive. The subjective approach can be performed quickly, with *all* the images read in a random manner in a single sitting that will minimize systematic errors. The most appropriate method thus depends on the specific scientific task at hand, with the subjective approach useful for performing pilot studies and the objective method ideal for obtaining definitive scientific results.

Acknowledgement. This work was supported in part by a US Army Grant No. DAMD 17-00-1003375. AE Burgess provided invaluable assistance in the development of the four AFC method.

References

1. Huda W, Sajewicz A, Ogden KM, and Scalzetti EM. "How good is the ACR accreditation phantom for measuring image quality in digital mammography?". *Academic Radiology* (2002) *(In press)*

2. Sajewicz A, Huda W, Hseuh D, Ogden KM, Scalzetti EM. "Observer performance and radiographic technique factors in digital mammography." SPIE Medical Imaging (2002) *(In press)*

3. Burgess AE. Comparison of receiver operating characteristic and forced choice observer performance measurement methods. *Medical Physics* 22 (1995) 643-655.

Experimental investigation of the dose and image quality characteristics of a digital mammography imaging system.

Walter Huda PhD[1]

Anthony M Sajewicz MD[1]

Kent M Ogden PhD[1]

David R Dance PhD[2]

[1]**Department of Radiology**

SUNY Upstate Medical University

750 E Adams Street

Syracuse, NY 13210

[2]**Department of Medical Physics**

The Royal Marsden NHS Trust

London SW3 6JJ, UK

ABSTRACT

The purpose of this study was to investigate the image quality and absorbed dose characteristics of a commercial digital mammography imaging system, and to identify an optimal x-ray tube voltage for imaging simulated masses in an average size breast with 50% glandularity. Images were taken of an ACR accreditation phantom using a LORAD digital mammography system with a Mo target and Mo filter. In one experiment, exposures were performed at 80 mAs with x-ray tube voltages varying between 24 and 34 kVp. In a second experiment, the x-ray tube voltage was kept constant at 28 kVp and the technique factor was varied between 5 and 500 mAs. The average glandular dose at each x-ray tube voltage was determined from measurements of entrance skin exposure and x-ray beam half value layer. Image contrast was measured as the fractional digital signal intensity difference for the image of a 4 mm thick acrylic disk. Image noise was obtained from the standard deviation in a uniformly exposed region of interest expressed as a fraction of the background intensity. The measured digital signal intensity was proportional to the mAs and to the $kVp^{5.8}$. Image contrast was independent of mAs, and dropped by 21% when the x-ray tube voltage increased from 24 to 34 kVp. At a constant x-ray tube voltage, image noise was shown to be approximately proportional to $(mAs)^{-0.5}$, which permits image contrast to noise ratio (CNR) to be modified by changing the mAs. At 80 mAs, increasing the x-ray tube voltage from 24 to 34 kVp increased the CNR by 78%, and increased the average glandular dose by 285%. At a constant lesion CNR, the lowest average glandular dose value occurred at 27.3 kVp. Increasing or decreasing the x-ray tube voltage by 2.3 kVp from the optimum kVp increased the average glandular dose values by 5%. These results show that imaging simulated masses in a 4.2 cm

compressed breast at ~27 kVp with Mo/Mo target/filter results in the lowest average glandular dose.

INTRODUCTION

The goal of mammography is to achieve the image quality required for a given detection task, whilst ensuring that the patient absorbed dose is kept as low as reasonably achievable.[1] In comparison to conventional screen-film imaging, the amount of radiation used to generate a digital image could be increased (or decreased) by over an order of magnitude with no significant change on the displayed image intensity. In addition, the quality of the x-ray beam (i.e., half value layer) used to acquire the digital radiograph may be adjusted by modification of the x-ray tube voltage (i.e., kVp).[2] It is of interest to quantify how modification of the x-ray tube mAs and kVp affect image contrast and noise, since this knowledge may be used to help optimize imaging performance.[3,4,5]

Choice of x-ray tube voltage and mAs will also affect the patient average glandular dose.[6] One important goal for using a digital imaging system is to attempt to keep patient doses as low as reasonably achievable.[7] In principle, this may be achieved by adjusting the radiographic technique factors (mAs and kVp) to maintain a constant image quality and selecting that technique factor which minimizes the patient dose. Information as to how image contrast to noise ratio (CNR) and patient dose vary with technique factors is of obvious importance. Knowledge of the dose versus image quality relationship will enable doses to be minimized at a constant image quality, or would permit any improvements in CNR to be quantitatively balanced by any corresponding increases in patient dose.[8,9,10,11]

Digital mammography separates the process of image acquisition from any subsequent image display, which should permit all the acquired image information to be optimally displayed to the observer and ensure that imaging performance is only limited by the

acquired CNR.[2,12] In this study, we investigated the CNR of a simulated mass and the corresponding absorbed dose performance of a commercial digital mammography system. Both the x-ray tube output (mAs) and x-ray tube voltage (kVp) were systematically varied, and the corresponding changes in image quality and dose were quantified. Results obtained in this study quantify the trade-offs between dose and image quality in digital mammography for the detection of simulated masses in an average size breast. Information obtained in this study is expected to help the process of optimizing clinical mammography.[13,14,15]

METHOD

Digital Mammography system

The full field of view digital mammography system (LORAD, Danbury CT) is a mosaic of twelve 1600 x 1600 pixel Charge Coupled Devices (CCDs) coupled by 2:1 fiber optic tapers to a large area Thallium activated Cesium Iodide (CsI:Tl) scintillator plate. The active image area of the image receptor covers an 18.6 cm x 24.8 cm field. The corresponding image pixel matrix size is 4800 by 6400. The pixel size at the scintillator surface is 40 μm, resulting in a Nyquist spatial frequency of 12.5 cycles per millimeter. A conventional linear grid (5:1 grid ratio) is employed to reject scattered x-rays.

A CsI:Tl scintillator converts the incident x-ray photons to light which is transmitted through the fiber optic tapers to the solid state CCD device. The CCD converts the visible photons to electrons, and the CCD output is digitized at 14-bit depth to produce the high dynamic range required for digital mammography.[16] The CsI:Tl scintillator-fiber optic-taper-CCD assembly is housed in a sealed chamber with CCDs being thermally stabilized at a low temperature to reduce system noise.

In screen-film radiography, dense objects appear white since little radiation is transmitted, which is the reverse of digital radiography where regions receiving the largest radiation exposure would appear the brightest. This digital mammography unit acquired image data with intensity values ranging from 0 to 16,383. The digital mammography system automatically inverts the gray scale values by subtracting the measured intensity from 16,383. The pixel values generated were corrected by subtracting them from 16,383, and they therefore correspond to the magnitude of the signal generated by the incident x-ray beam intensity.

Exposure of ACR phantom

A standard American College of Radiology (ACR) phantom[17] was used to acquire digital images at different values of x-ray tube voltage (kVp) and tube current-exposure time product (mAs). The phantom has a composition and a thickness that is equivalent to a 4.2 cm compressed breast consisting of 50% glandular and 50% adipose tissue. Figure 1 shows a representative image obtained of the accreditation phantom at 28 kVp and 80 mAs, showing the fibers, microcalcification specks and masses. Also depicted in Figure 1 is an acrylic disk (4 mm thick and 1 cm diameter) located in the bottom left hand corner, above the bottom row of masses. Detection of this disk was the diagnostic task used in this study to quantify how image quality of this digital mammography system varied with changes in radiographic technique.

The x-ray spectrum was generated using a Molybdenum target and a Molybdenum filter (25 μm). In one experiment, the x-ray tube voltage was kept constant at 28 kVp and digital images were generated at tube current-exposure time product values ranging from 5 to 500 mAs. In a second experiment, the tube current-exposure time product was kept constant at 80 mAs, and the x-ray tube voltage was varied between 24 and 34 kVp. In addition, a series of five additional repeat images were obtained at 28 kVp and 80 mAs to provide data on the experimental precision of the image quality measurements. Table 1 summarizes the three series of experiments performed with the ACR accreditation phantom. Measurements were made of the entrance skin exposure and half value layer using the recommended protocols of the ACR. Entrance skin exposure measurements were converted into corresponding values of average glandular dose for a standard 4.2 cm compressed breast using data provided in the ACR manual.[16]

Contrast & noise

The ACR accreditation phantom image was imaged with an added disk that is 4mm thick and 1cm in diameter as depicted in Figure 1. Relative values of disk image contrast (C) were obtained as the difference between the average disk intensity (I_{disk}) and the surrounding average background intensity ($I_{background}$), and normalized by the average background intensity, so that

$$C = (I_{background} - I_{disk})/ I_{background} \quad (1).$$

The value of C in equation (1) was always a positive value since the intensity in the background region was greater than that behind the disk. The region of interest (ROI) used to determine the average signal intensities in the background and disk regions was a square with a size of approximately 55 x 55 pixels. The ROI was located at the center of the disk to determine the value of I_{disk}, and 5 mm below the disk shown in Figure 1 for the determination of $I_{background}$.

In the background area with a nominal uniform exposure, the mean intensity value is $I_{background}$, and the measured standard deviation is σ. The relative noise level, N, is then given by

$$N = \sigma/I_{background} \quad (2).$$

The contrast to noise ratio (CNR) was obtained from the ratio of measured contrast (equation (1)) to the corresponding noise (equation (2)). The CNR is thus given by

$$CNR = (I_{background} - I_{disk})/\sigma \qquad (3).$$

The CNR is the ratio of the image contrast to the random fluctuations about background intensity value measured using the same scale. Equation (3) is independent of the lesion disk diameter, and does not predict imaging performance for the "detection" of this type of disk in a uniform background. Only relative changes of CNR are used in this study, and no significance is attached to specific values of the CNR defined by equation (3) and reported here.

RESULTS

Digital detector characteristics

Figure 2 (a) shows the average signal intensity in the background region plotted as a function of selected mAs value at a constant x-ray tube voltage (28 kVp). Figure 2 (a) shows the expected linear response, with a slope of about 30 pixel values per unit mAs. It is also evident that the digital system has not saturated at the maximum 500 mAs value used in this experiment; extrapolation of the data in Figure 2 (a) shows that the system would saturate at a tube current-exposure time product of ~540 mAs for an x-ray tube voltage of 28 kVp.

Figure 2 (b) shows the average background signal intensity as a function of kVp at a constant tube current-exposure time product (80 mAs), where both the ordinate and abscissa are presented using a logarithmic scale. This shows the supra-linear response expected when the x-ray tube voltage is increased at a constant mAs value. The solid line depicted in Figure 2 (b) has a slope of 5.80, and thus the measured signal intensity is proportional to $kVp^{(5.8)}$.

Experimental precision

Seven repeat experiments were available for analysis at 28 kVp and 80 mAs. The intensity values in the disk region ranged from 1883 to 1886, and the intensity values in the background region ranged from 2414 to 2418. The measured standard deviation in the disk region ranged from 16 to 17, and the measured standard deviation in the background region ranged from 18 to 19. These data clearly indicate that the digital mammography system is very stable. It is also evident that the precision of any noise measurements will be limited to only two significant figures.

The measured precision for image contrast was < 0.2%, and the corresponding precision for image noise was 2.9%. The overall measured precision for disk CNR was 3%. Error bars in the figures below indicate this experimental precision at data presented for 28 kVp and 80 mAs, unless the size of the error bar was too small to be visible.

Contrast and noise

As expected, image contrast was found to be independent of the selected mAs value. Figure 3 shows image contrast as a function of x-ray tube voltage, which exhibits the expected decrease in contrast with increasing x-ray tube voltage. Increasing the x-ray tube voltage from 24 to 34 kVp reduced image contrast by 21% (i.e., 1.9% per unit increase in kVp).

Figure 4 (a) shows how the image noise varied with mAs where the ordinate and abscissa are plotted on logarithmic scale. The solid line is a least squares fit of a straight line to the experimental data (r^2 = 0.98), with a slope of -0.506. Since the slope of the curve in Figure 4 (a) is very close to the value expected for an imaging system with a noise that is determined by quantum mottle (i.e., slope of -0.500), this digital mammography system may be taken to be quantum noise limited over the complete dynamic range investigated (i.e., 5 to 500 mAs). Figure 4 (b) shows the measured image noise versus x-ray tube voltage. The data in Figure 4 (b) show that as the x-ray tube voltage increases, the noise level is markedly reduced. Increasing the x-ray tube voltage from 24 to 34 kVp reduced the image noise by approximately 55.8%.

We investigated the importance of the location of the background ROI for determining image noise and contrast. A second ROI was identified 5 mm above the disk shown in

Figure 1, and we compared the measured value of contrast and noise with those described above for an ROI located 5 mm below the disk. For the 12 images in the mAs series, the average intensity ratio of two background regions was 1.002, and the corresponding average ratio of the measured standard deviations was 1.026. These data indicate that the choice of background ROI location had no significant effect on the resultant image noise and contrast values.

Figure 5 shows the CNR data for varying x-ray tube voltage at a constant 80 mAs. Raising the x-ray tube voltage from 24 to 34 kVp increased the CNR by 78%. Increasing the kVp reduces image contrast (see Figure 3), but this is more than offset by a corresponding reduction in image noise (see Figure 4 (b)). The rate of increase of CNR with x-ray tube voltage falls off with increasing kVp. At 24 kVp, the value of CNR increases by 14% per kVp, at 28 kVp the rate of increase falls to 6.4% per kVp, and at 34 kVp the CNR increases by only 0.8% per kVp.

Radiation dose

Table 2 summarizes the absorbed dose data obtained for this digital mammography system as a function of x-ray tube voltage. At 28 kVp and 80 mAs, the average glandular dose was 2.16 mGy. At this constant x-ray tube voltage, the average glandular dose is directly proportional to the selected mAs value. At a constant 80 mAs, increasing the x-ray tube voltage from 24 to 34 kVp increased the average glandular dose from 1.12 to 4.32 mGy (i.e., 285%).

For a given x-ray tube voltage, the image CNR can be adjusted by modification of the mAs used to acquire these images. Figure 6 shows how the mAs would need to be reduced with increasing x-ray tube voltage to maintain the CNR observed at 24 kVp.

Figure 7 shows the variation of the average glandular dose with x-ray tube voltage at a constant CNR for the detection of this type of simulated mass lesion. For this standard 4.2 cm compressed breast with a 50% glandularity, the lowest radiation exposure occurs at 27.3 kVp when image quality (i.e., CNR) is kept constant. Increasing or decreasing the x-ray tube voltage by 2.3 kVp from the optimum kVp increased the average glandular dose values by 5%.

DISCUSSION

The data in Table 2 indicate that the x-ray tube output (i.e., entrance skin exposure) in mammography varies by approximately $kVp^{3.05}$. By comparison, a kVp^2 dependence is normally expected in the diagnostic imaging range.[18] For an average size breast, the detected intensity shows an even greater dependence on x-ray tube voltage (i.e., $kVp^{5.8}$), that reflects the non-linear dependence in x-ray beam transmission through the ACR phantom as a function of the x-ray tube voltage. These data demonstrate that small changes in x-ray tube voltage will have relatively large effects on the x-ray tube output and detected signal intensities. Changing the x-ray tube voltage from 28 to 29 kVp, for example, increased x-ray tube output by ~11%, and the corresponding detected signal intensity by ~22%.

The slope of the curve in Figure 4 (a) is approximately -0.5 demonstrating that quantum mottle is the dominant source of image noise. However, the experimental data shown in Figure 4 (a) deviate from a simple power law relationship with an exponent of -0.5. It is evident that there are additional noise sources in this digital mammography imaging system. Electronic noise is the most likely additional noise source at low exposures; structured noise and a non-linear response of the CCD are the most likely noise sources at the highest exposure levels[i]. Nonetheless, quantum mottle is the dominant source of image noise in the clinically relevant exposure range taken to be between 40 and 200 mAs; this feature permits image CNR to be readily adjusted by modifying the selected mAs. When performing clinical mammography, increasing the mAs by a factor of two is expected to improve the image CNR by approximately 41%.

[i] Private communication Dr Z Jing

The detected x-ray signal varies as $kVp^{5.8}$, and if this were simply due to a proportional increase in the number of photons, the slope of a plot of log (noise) versus log (kVp) would have a slope of -2.9 as shown by the dashed line in Figure 4 (b). The experimental data deviate significantly from this value because the increased signal is a result of increased energy deposition in the CsI detector due to the higher energy photons transmitted through the phantom at higher x-ray tube voltages. Increasing the x-ray tube voltage from 28 to 34 kVp reduced the noise by 32%, whereas a slope of -2.9 would have produced a reduction of 44%.

For screen-film mammography, current regulations in the United States limit the average glandular dose to 3 mGy, and typical clinical systems normally operate at average glandular doses of about 1.5 mGy.[19] The average glandular dose at 28 kVp/80 mAs on this mammography imaging system was 2.16 mGy. At 28 kVp, using 56 mAs would result in patient doses comparable to those encountered in screen-film radiography (i.e., 1.5mGy), whereas reducing the x-ray tube voltage to 25 kVp would require approximately 90 mAs.

The choice of x-ray tube voltage in screen-film radiography is guided by an attempt to maximize image contrast. In digital mammography, however, selecting the x-ray tube voltage and mAs should achieve a signal-to-noise ratio that enables an accurate diagnosis to be made, and which also minimizes the patient dose.[20] The data in Figure 7 show that for the task of detecting a simple disk type lesion, 27.3 kVp results in the lowest average glandular dose, and would therefore be deemed to be the optimal x-ray tube voltage. It is possible to define image noise (see equation (2)) as the relative standard deviation for an ROI located in the disk rather than the background region. An analysis of the relative

CNR vs kVp with the noise defined in this alternative manner resulted in an optimum kVp of 27.8 kVp, which differs by 0.5 kVp from the value obtained when the noise was defined using equation (2).

The experimental results obtained in this study can be compared with recent calculations performed by Dance *et al*[21] performed with a Gd_2O_2S screen. For a Molybdenum target and Molybdenum filter similar to those used in this study, Dance *et al* observed a radiation dose minimum at 26.3 kVp for a 5 mm thick glandular tissue lesion in a 4 cm thick breast with 50% glandularity. The optimum x-ray tube voltage for a mass was close to the dose minimum of 27.0 kVp obtained for a 200 μm calcification. It is noteworthy that these theoretical calculations also showed that x-ray tube target/filter combinations that increased the x-ray photon energy (e.g., Mo/Rh; Rh/Al; Rh/Rh; W/Rh) could reduce patient doses by up to 15% whilst maintaining a constant level of image quality.

In digital mammography, imaging performance is task dependent[22] and will generally be different for microcalcifications and masses.[23,24] The photon energy dependence of lesion detection will depend on the type of object that is being detected. Accordingly, there may be different optimal values for malignant masses and calcifications because of the different effective atomic numbers of these types of materials.[25] Detection performance may also depend on the specific size and shape of the lesion, breast composition and thickness,[26] as well as the nature of the structured breast background.[27] In these cases, a detailed analysis of the spatial frequency dependent noise and resolution performance of the mammography imaging system may be required to generate a full description of the overall signal to noise ratio. It is possible that for more complex imaging tasks than the

one adopted in this study could result in optimal x-ray tube voltages that differ from the value of 27.3 kVp.

Digital mammography systems are likely to significantly differ in terms of the x-ray spectra[28], and also use different types of x-ray detector systems to acquire the image. The object under investigation was relatively large, and thus spatial resolution is not a significant factor to be included in analyzing relative imaging performance with radiographic technique factors. Differences between this imaging system and other comparable types of digital mammography systems relate to the effective photon energy of the x-ray beam as well as the scatter to primary ratio in the detected x-ray signal. Differences in effective photon energy and scatter to primary ratio at the image receptor could result in different values of the optimum x-ray tube voltage for this type of imaging task. One important advantage of using a standard phantom for assessing dose and image quality is the ability to directly compare two systems.[29] The results reported in this study were obtained with an ACR phantom readily available in other laboratories that permits our results to be directly inter-compared with those achievable for any other type of digital mammography imaging system.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. Zhenxue Jing PhD for assistance with the experimental work and useful discussions on digital mammography. LORAD provided access to the digital mammography imaging system and the ACR accreditation phantom. This work was supported in part by a US Army Grant No. DAMD 17-00-1003375.

TABLES

Table 1. Summary of digital radiographs obtained of the ACR phantom

Series	Constant parameter	Variable parameter	# of images
1	mAs (80)	kVp (24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34)	11
2	kVp (28)	mAs (5, 10, 20, 40, 80, 120, 160, 240, 325, 400, 450, 500)	12
3	mAs (80) & kVp (28)	Five repeat examinations (obtained to estimate the experimental precision)	5*

*A *total* of seven images were available for the precision measurements, which included exposures at 28 kVp and 80 mAs in series 1 and 2.

Table 2. Absorbed dose summary for the digital mammography system obtained at a constant tube current-exposure time value (80 mAs).

X-ray tube voltage (kVp)	Entrance skin exposure (R*)	Half value layer (mm Al)	Mean glandular dose (mGy)
24	0.718	0.303	1.12
25	0.841	0.316	1.35
26	0.960	0.330	1.61
27	1.08	0.340	1.88
28	1.22	0.350	2.16
29	1.35	0.360	2.48
30	1.49	0.369	2.82
31	1.63	0.376	3.14
32	1.79	0.384	3.50
33	1.94	0.389	3.88
34	2.11	0.400	4.32

*1 R = 2.58×10^{-4} C kg^{-1}

REFERENCES

[1] A. G. Haus and M. J. Yaffe, in Physical Aspects of Breast Imaging-Current and Future Considerations, RSNA 1999.

[2] M. J. Yaffe, in *Digital Mammography*, proceedings of the RSNA Categorical Course in Breast Imaging, Chicago, Illinois, (1999), p. 229-238.

[3] R. J. Jennings, R. J. Eastgate, M. P. Siedband and D. L. Ergun, "Optimal x-ray spectra for screen-film mammography," Med. Phys. **8**(5), 629-39 (1981).

[4] R. Fahrig and M. J. Yaffe, "Optimization of spectral shape in digital mammography: Dependence on anode material, breast thickness, and lesion type," Med. Phys. **21**(9), 1473-81 (1994).

[5] R. E. Hendrick and E. A. Berns, in Optimizing Mammographic Techniques, proceedings of the RSNA Categorical Course in Breast Imaging, Chicago, Illinois, (1999), p. 79-89.

[6] L. N. Rothenberg, in Exposures and Doses in Mammography, proceedings of the RSNA Categorical Course in Breast Imaging, Chicago, Illinois, (1999) p.91-97.

[7] Protection of the Patient in Diagnostic Radiology, ICRP Publication 34, 1982.

[8] L. Stanton *et al.*, "Screen-Film Mammographic Technique for Breast Cancer Screening," Radiology **163**, 471-479 (1987).

[9] L. Desponds *et al.*, "Influence of anode and filter material on image quality and glandular dose for screen-film mammography," Phys. in Med. & Biology **36**(9), 1165-82 (1991).

[10] E. L. Gingold, X. Wu and G. Barnes, "Contrast and Dose with Mo-Mo, Mo-Rh, and Rh-Rh Target Filter Combinations in Mammography," Radiology **195**, 639-44 (1995).

[11] K. C. Young, M. L. Ramsdale and A. Rust, "Dose and image quality in mammography with an automatic beam quality system," British J. of Radiology **69**(822), 555-62 (1996).

[12] E.D. Pisano, "Current status of full-field digital mammography," Radiology **214**, 26-29 (2000).

[13] C. Kimme-Smith *et al.*, "Mammograms obtained with rhodium vs molybdenum anodes: contrast and dose differences," Am. J. Roentgenology **162**(6), 1313-7 (1994).

[14] K. C. Young, M. G. Wallis and M. L. Ramsdale, "Mammographic Film Density and Detection of Small Breast Cancers," Clinical Radiology **49**, 461-5 (1994).

[15] A. C. Thilander-Klang *et al.*, "Influence of Anode-Filter Combinations on Image Quality and Radiation Dose in 965 Women Undergoing Mammography," Radiology **203**, 348-54 (1997).

[16] A. Maidment, R. Fahrig and M. J. Yaffe, "Dynamic range requirements in digital mammography," Med. Phys. **20**(6), 1621-34 (1993).

[17] American College of Radiology (ACR) Mammography Quality Control Manual, ACR Reston VA (1999).

[18] J. T. Bushberg, J. A. Seibert, E. M. Leidholdt and J. M. Boone, *The Essential Physics of Medical Imaging*, 3rd ed. (Williams & Wilkins, Baltimore, 1994), p. 99.

[19] W. Huda, T. LaVoy and K. Ogden, "Radiographic techniques in screen-film mammography," Med. Phys. **28**, 1187 (2001).

[20] A. G. Haus and M. J. Yaffe, "Screen-film and digital mammography. Image quality and radiation dose considerations," Radiologic Clinics of North America **38,** 871-898 (2000).

[21] D. R. Dance et al., "Influence of anode/filter material and tube potential on contrast, signal-to-noise ratio and average absorbed dose in mammography: a Monte Carlo study," British J. Radiology **73**, 1056-67 (2000).

[22] "Medical Imaging-The Assessment of Image Quality," International Commission on Radiation Units and Measurements Report **54**, 1996.

[23] W. F. Good et al., "Detection of masses and clustered microcalcifications on data compressed mammograms: an observer performance study," AJR **175**, 1573-76 (2000).

[24] E. D. Pisano et al., "Radiologists' preferences for digital mammographic display," Radiology **216**, 820-830 (2000).

[25] W. Huda, A. Krol, Z. Jing and J. M. Boone, "Signal to noise ratio and radiation dose as a function of photon energy in mammography," SPIE Medical Imaging **3336,** 355-63 (1998).

[26] B. J. McParland and M. M. Boyd, "A comparison of fixed and variable kVp technique protocols for film-screen mammography," Br. J. Radiol. **73**, 613-626 (2000).

[27] A. E. Burgess, F. L. Jacobsen and P. F. Judy, "Human observer detection experiments with mammograms and power-law noise". Medical Physics **28,** (2001) 419-437.

[28] C. Kimme-Smith, "New digital mammography systems may require different x-ray spectra and, therefore, more general normalized glandular dose values," Radiology **213**, 7-10 (1999).

[29] P. E. Undrill, A. D. O'Kane and F. J. Gilbert, "A comparison of digital and screen-film mammography using quality control phantoms," Clinical Radiology **55**, 782-790 (2000).

List of Figure Captions

Figure 1. Digital radiograph of the ACR accreditation phantom obtained at 28 kVp and 80 mAs. The white circle in the bottom left is due to the presence of the 4 mm thick disk (1 cm diameter) added to the phantom for this exposure.

Figure 2. Plot of the background disk intensity versus selected radiographic technique: a) intensity versus mAs, where the solid line is a least squares fit to straight line ($r^2 > 0.99$); b) intensity versus x-ray tube potential where both abscissa and ordinate are on a logarithmic scale and the solid line is a least squares fit to a straight line ($r^2 > 0.99$).

Figure 3. Plot of image contrast versus x-ray tube potential; the solid line is a least squares fit to straight line for the experimental data points ($r^2 > 0.99$).

Figure 4. Plot of noise versus selected radiographic technique: a) noise versus mAs, where both abscissa and ordinate are on a logarithmic scale and the dotted line is a least squares fit to a straight line ($r^2 = 0.98$); b) noise versus x-ray tube potential, where both abscissa and ordinate are on a logarithmic scale. The dashed line in b) has been drawn with a slope of -2.9 (see text for discussion).

Figure 5. Plot of contrast to noise ratio versus x-ray tube potential at a constant 80 mAs. Dotted line is a least squares fit to a second order polynomial ($r^2 = 0.99$).

Figure 6. Plot of the mAs reduction factor required to maintain the CNR obtained at 24 kVp, where the solid line is a least squares fit to a fourth order polynomial ($r^2 > 0.99$).

Figure 7. Plot of the average glandular dose as a function of x-ray tube potential obtained at a constant contrast to noise ratio. The solid line is a least squares fit to a fourth order polynomial ($r^2 > 0.99$).

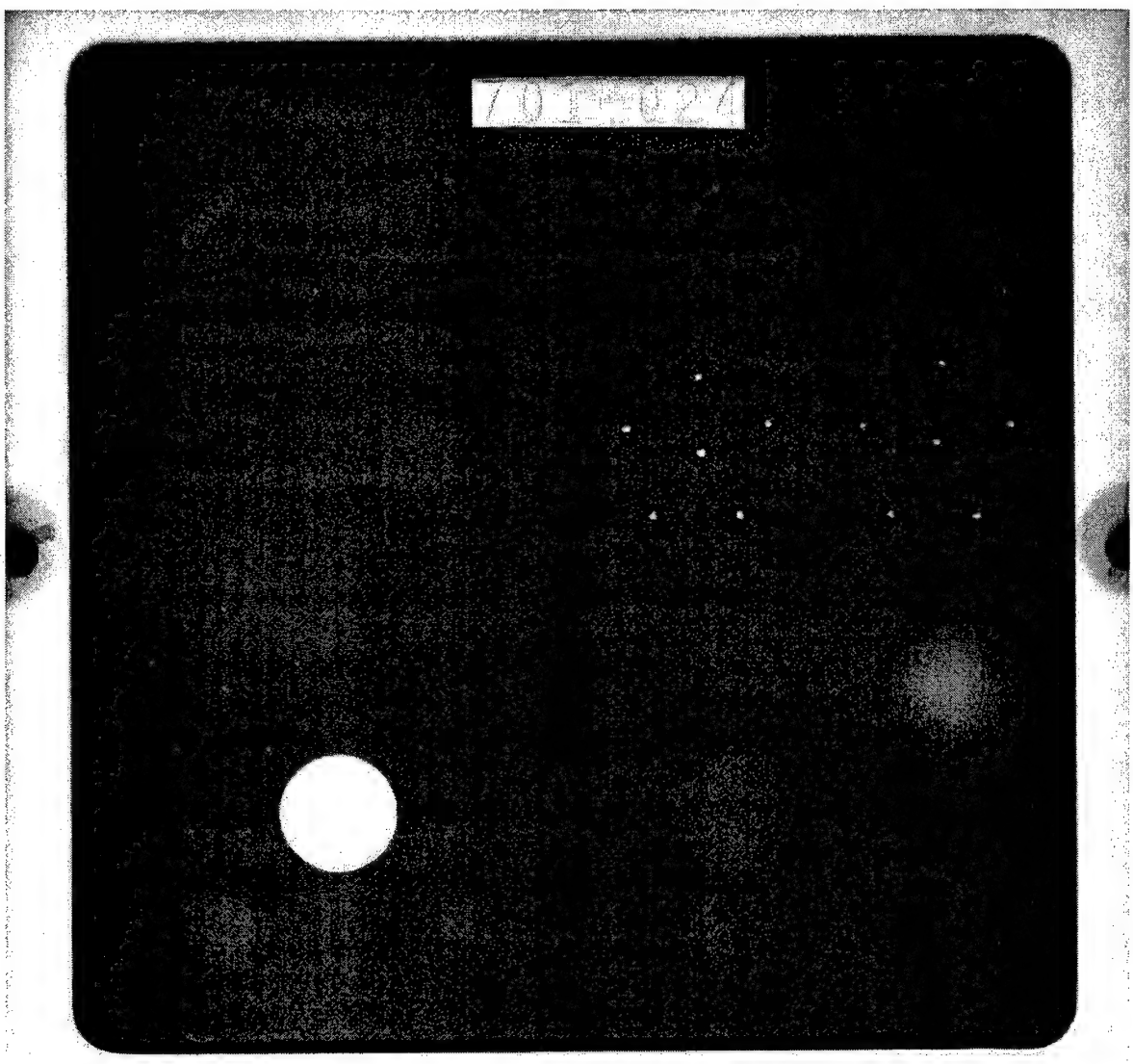
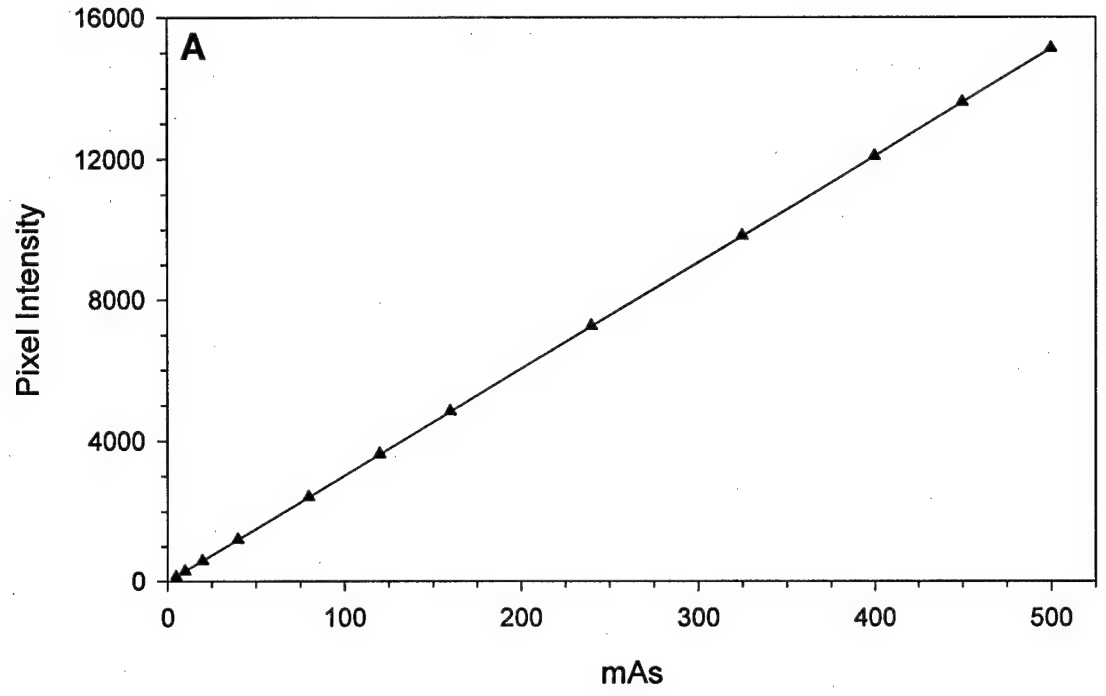


FIGURE 1


A
16000
12000
8000
4000
0
Pixel Intensity
0
100
200
300
400
500
mAs

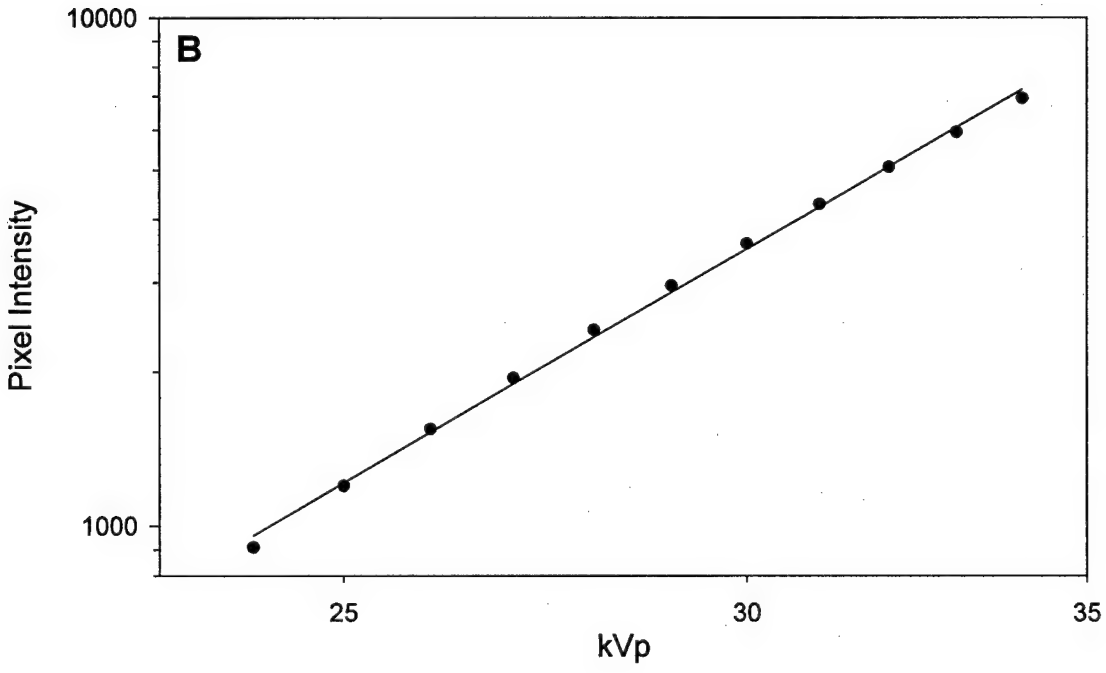

B
10000
1000
Pixel Intensity
25
30
35
kVp

FIGURE 2

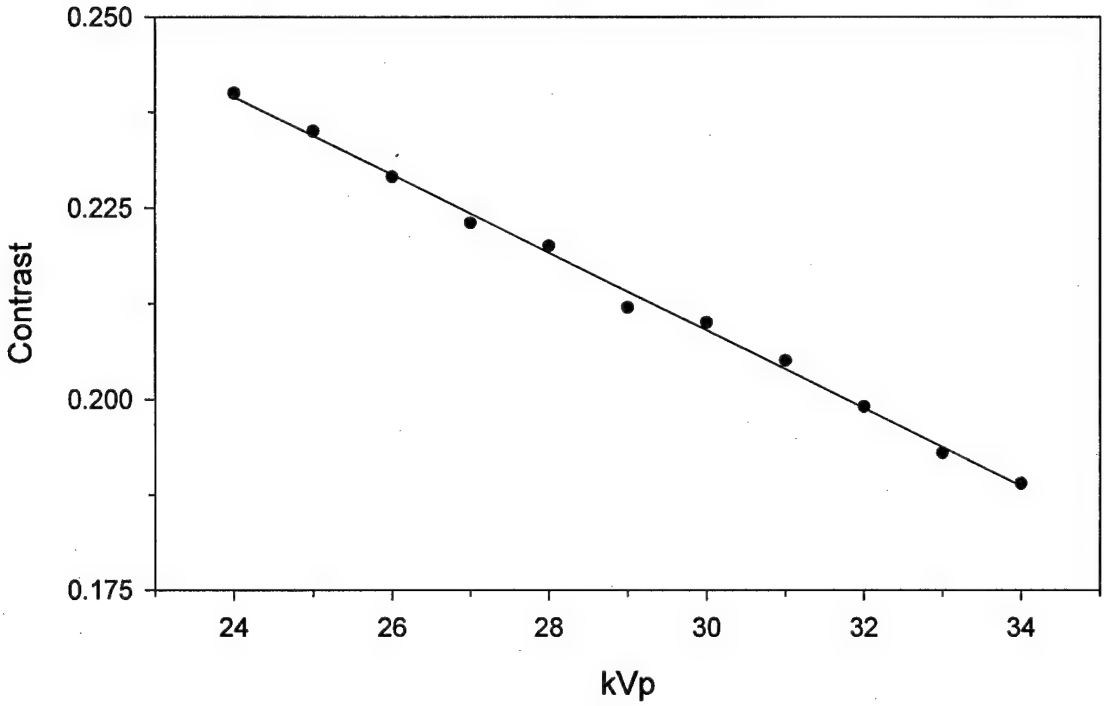
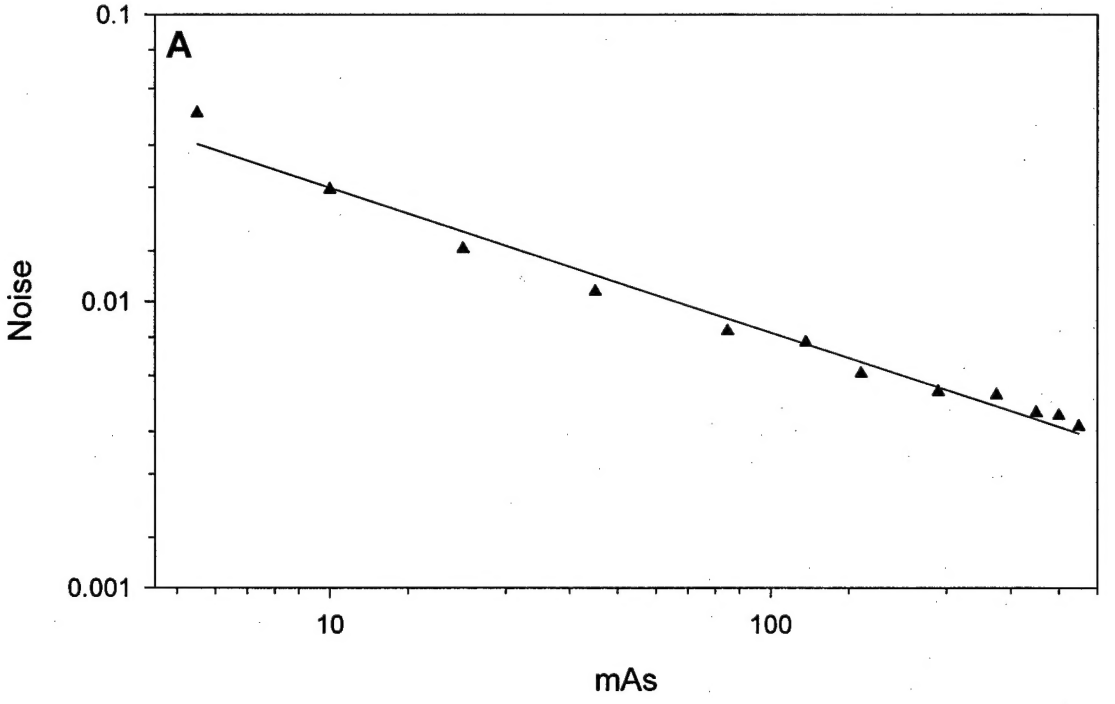


FIGURE 3


A
0.1
0.01
0.001
Noise
10
100
mAs

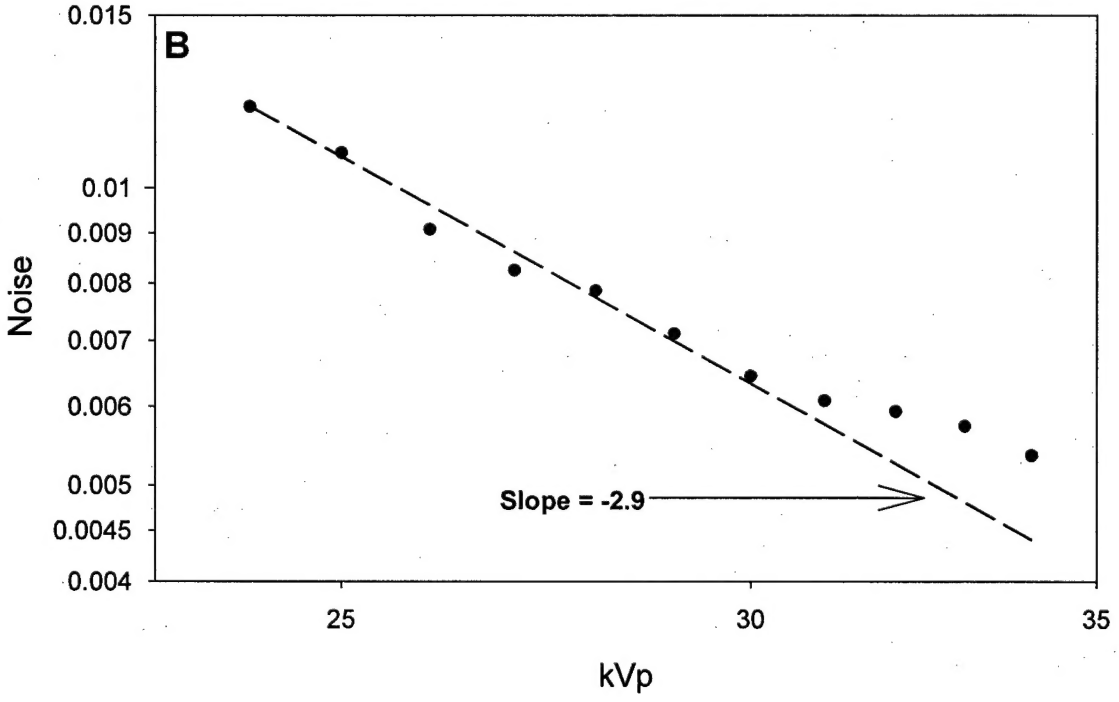

B
0.015
0.01
0.009
0.008
0.007
0.006
0.005
0.0045
0.004
Noise
Slope = -2.9
25
30
35
kVp

FIGURE 4

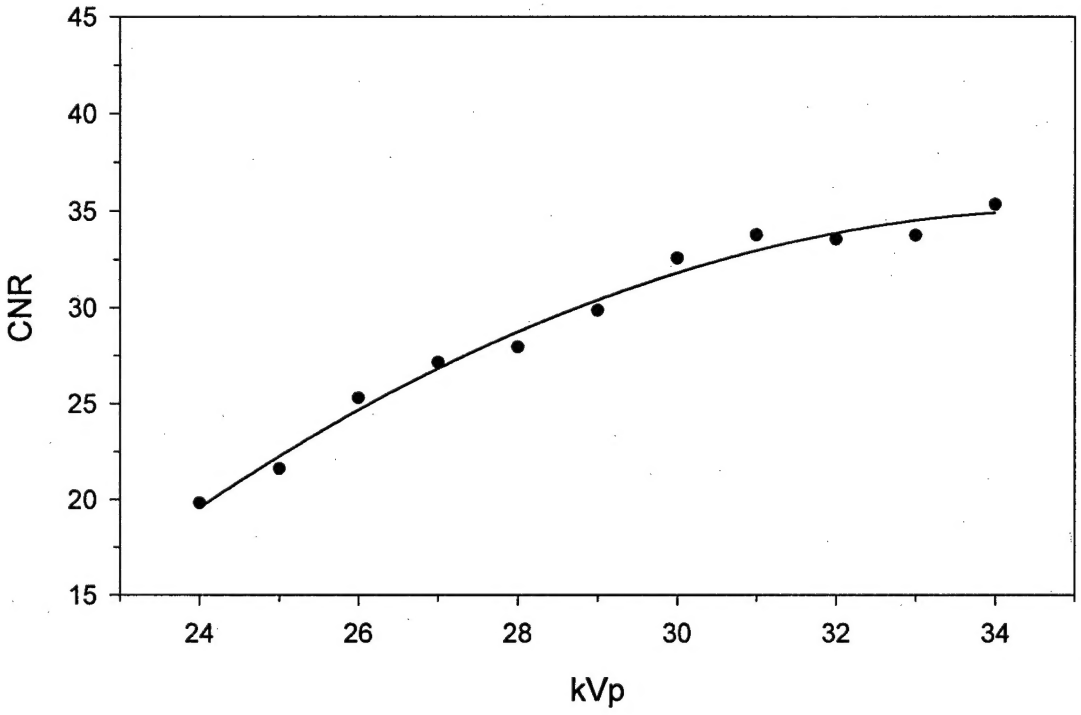


FIGURE 5

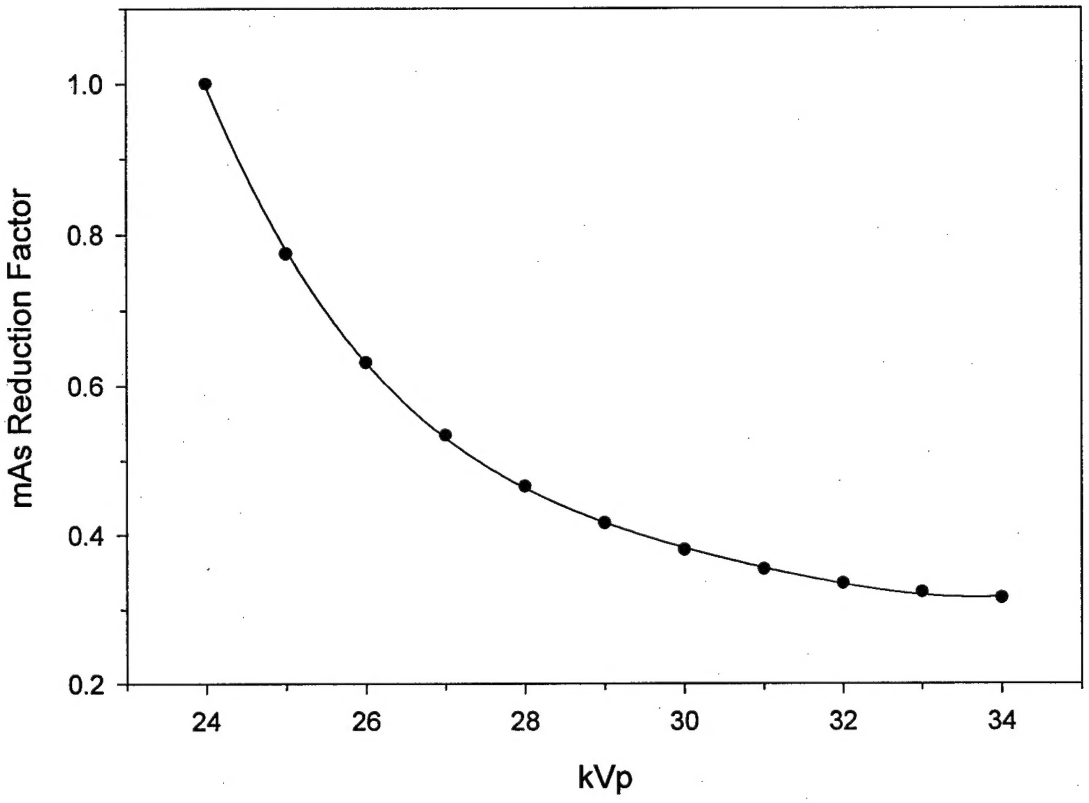


FIGURE 6

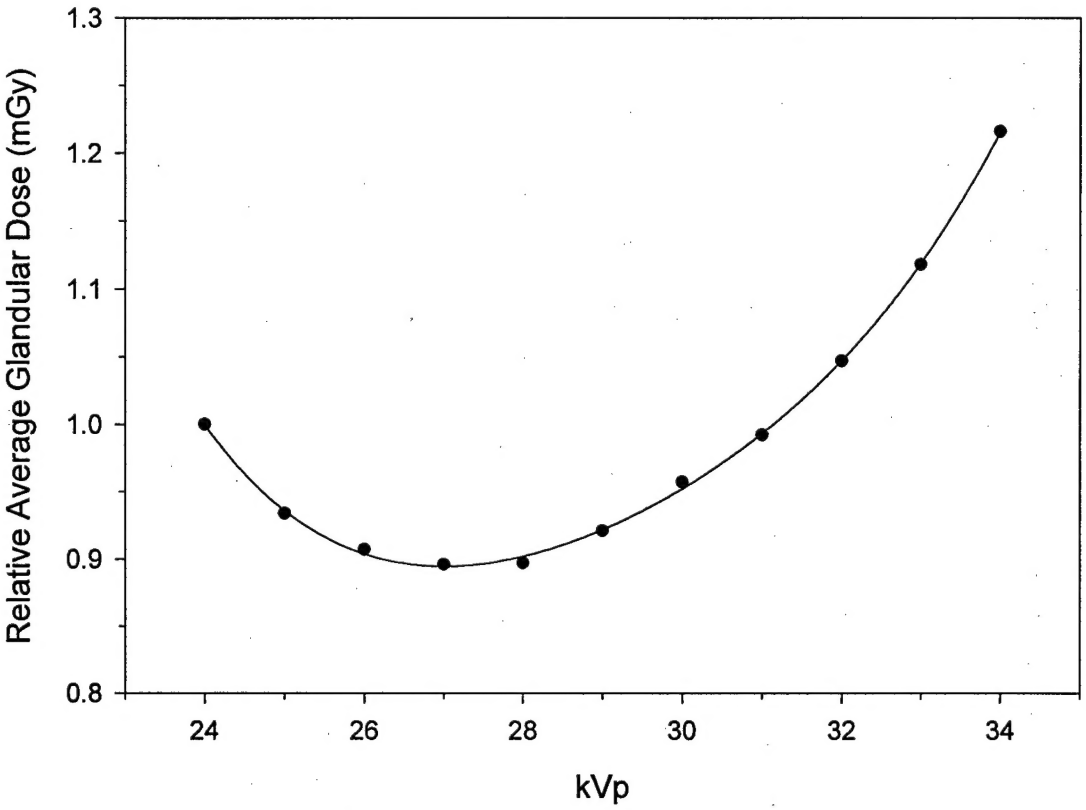

Relative Average Glandular Dose (mGy)
1.3
1.2
1.1
1.0
0.9
0.8
24
26
28
30
32
34
kVp

FIGURE 7